CONFOCAL SPECTROSCOPY IN LIVING CELLS
CHROMATIN AND PROTEIN INTERACTIONS OF HIV-1 INTEGRASE CO-FACTOR LEDGF/p75

Jelle HENDRIX

Members of the Examination Committee
prof. dr. Yves Engelborghs, promotor
prof. dr. Zeger Debyser, promotor
prof. dr. Johan Hofkens, chairman
prof. dr. Johan Robben, secretary
prof. dr. Carmen Bartic
prof. dr. Don C. Lamb (University of Munich)
dr. Jan De Rijck

Dissertation presented in partial fulfillment of the requirements for the degree of Doctor in Science

March 2010

ISBN 978-90-8649-323-4
D/2010/10.705/16

Dankwoord

Eens alles netjes neergeschreven staat in een boekje, lijkt alsof het behalen van een doctoraat voor de hand liggend is, na je hogere studies gewoon 4 jaar onderzoek doen op een bepaald wetenschappelijk thema, al je resultaten neerschrijven en *klaar is kees*. Aan een doctoraat werken is véél meer dan dat. Er komen ontelbare uitdagingen, ups en downs en ook wel wat stress bij kijken. Gelukkig zijn er heel wat mensen geweest die mij gedurende deze periode hebben geholpen. Zonder hen zou dit boekje nooit tot stand zijn gekomen, daarom wil ik heb in dit dankwoord even vermelden.

In de eerste plaats is er mijn promotor, prof. Yves Engelborghs. Bedankt dat U mij altijd in alle vrijheid mijn onderzoek heeft laten uitvoeren, dat U er altijd was als ik mijn resultaten (of nood daaraan) met U wou bespreken en dat U me heeft geleerd wat wetenschappelijke zelfstandigheid betekent. Bovenal wil ik U bedanken omdat U altijd objectief en eerlijk was, en dat U ervoor zorgde dat ik steeds met mijn beide *voetjes* op de grond bleef staan. Tot slot, de vele conferenties waar ik samen met U naartoe ging waren een geschenk uit de hemel, onder andere omdat ze mijn kijk op de biofysische wereld beduidend hebben verbreed.

Een gedeelde eerste plaats is er voor mijn andere promotor, prof. Zeger Debyser. Onze samenwerking was van onschatbare waarde voor mijn wetenschappelijke ontplooiing. Het (vaak van op een afstand) meedraaien in Uw groep heeft me geleerd om waardevolle biologische vragen te formuleren, om elk biofysisch experiment uit te voeren met de biologische relevantie ervan in het achterhoofd en om mijn onderzoek efficiënt te plannen. Bedankt hiervoor, maar ook voor uw enthousiasme en positieve houding, iets wat me hielp als ik door het bos de bomen niet meer zag.

I would like to thank the members of the examination committee, prof. Lamb, prof. Bartic, prof. Hofkens, prof. Robben and dr. De Rijck for the critical evaluation of my thesis.

In de BioMol groep bedank ik de mensen die op één of andere (wetenschappelijke, persoonlijke, ontspannende) manier hebben bijgedragen aan mijn onderzoek: Abel, Annelies,

Arnout, Bendy, Cicerone, Ellen, Els, Eric, Jessika, Jo, Karin, Marc, Melanie, Mieke, Mila, Renée, Rik S., Sangeeta, Sarah en Werner. De labmeetings, discussies en hulp bij experimenten waren natuurlijk zeer nuttig maar de aangename 'Heverlee' sfeer die er heerste op het labo, gecombineerd met de (talloze) labfeestjes, compenseerde voor het soms wel veeleisende onderzoek. Tot slot, ik weet dat ik velen nog een klein fortuin moet terugbetalen aan *colaatjes* uit de cola-automaat, ik zal dat ooit nog doen, beloofd!

In de MolMed groep bedank ik ook iedereen die op één of andere manier heeft bijgedragen aan mijn onderzoek. In het bijzonder wil ik Jan, Rik, Frauke en de hele HIV groep bedanken, op de labmeetings maakten jullie altijd tijd om mijn experimenten op een geïnteresseerde en constructieve manier te bediscussiëren, jullie positieve houding tegenover mijn werk heeft er toe bijgedragen dat ik altijd bleef geloven in mijn onderzoek. De verschillende feestcomité's dank ik voor de organisatie van spetterende kerstfeestjes, barbecue's en labweekends!

At the Laboratory for Photochemistry and Spectroscopy I would like to thank a couple of persons. Eerst en vooral, prof. Johan Hofkens, bedankt dat ik gedurende de voorbije jaren de faciliteiten in Uw labo heb mogen gebruiken, dat heeft mijn onderzoek én wetenschappelijke kennis en interesse werkelijk in de juiste richting doen evolueren. Cristina, you were a very helpful co-author on my first article, thank you very much. Jun-ichi and Peter, your doors were always open if I needed help with experiments, thank you both.

Op het Laboratorium voor Virologie en Chemotherapie in het Rega-instituut wens ik ook een aantal mensen te bedanken. Dirk, bedankt dat ik in alle vrijheid je microscoop mocht gebruiken. Els, bedankt om me tijdens mijn eerste ervaringen met de microscoop bij te staan met al mijn kleine en grote vragen.

Prof. Gielens, het Biochemie practicum heeft aan U een niet te onderschatten en onmisbare leiding en hulp, bedankt dat U zich altijd heeft ingezet daarvoor. Ook wil ik U bedanken dat ik refractometrie metingen heb mogen uitvoeren op uw labo.

Alle mensen van de Arenberg doctoraatsschool, het Geel Huis, de facilitaire en departementale diensten wens ik te bedanken voor hun bereidheid om mijn administratieve en technische probleempjes (zoals het zichzelf opsluiten in het labo) altijd snel en efficiënt op te lossen.

Ik dank ook de mensen van Carl Zeiss Benelux: Johannes Broeckhove, Wilfred Rijnsburger en Dirk van Meensel, omdat ze altijd vrijwel onmiddellijk te hulp sprongen wanneer de microscoop het (weer eens) liet afweten of wanneer ik (weer eens) een vraag had.

Ik dank het Instituut voor de Aanmoediging van Innovatie door Wetenschap en Technologie in Vlaanderen (IWT), voor de specialisatiebeurs die ze me toegekend hebben, en die me toeliet mijn onderzoek in alle vrijheid uit te voeren.

Beste ouders, vrienden en familie, ik denk niet dat iemand ooit dacht dat ik het zó lang zou trekken in Leuven, ik eerlijk gezegd ook niet. Menig ouder zou zich zorgen maken als hun zoon op zijn 26ste nog geen eigen auto of huis had, nog niet getrouwd was, nog geen kinderen in het vooruitzicht had, of een 'echte' job, nog geen belastingen betaalde en zelfs nog geen eigen wasmachine had… maar ik denk dat dit boekje dat wel allemaal kan compenseren, toch? Bedankt voor alles.

Tomas, een dankwoord schrijven zonder jou te vermelden, dat kan ik niet maken. Je hebt me vaak genoeg geholpen met redeneren over mijn resultaten, over autocorrelaties, over Matlab, over *weet-ik-veel* wat nog allemaal, of je was vaak gewoon een luisterend oor. Danku!

Liefste Sarah, bedankt dat je altijd (onder lichte dwang) naar mijn theorie, hypothese of verhaal wou luisteren. Of ik ooit *Fluorescence Correlation and Coloration Spectroscopy* metingen heb uitgevoerd op de *Latent Enzyme Deficiency Growth Factor* weet ik echter niet, maar dat doet er niet toe, toch? Bedankt dat je mij altijd probeerde te kalmeren wanneer ik weer eens aan het *flippen* was omdat ik weer eens te laat was met mijn powerpoint-presentatie, dat je 's avonds na de dansles soms nog meeging naar het labo, dat je mij altijd terug moed insprak wanneer de experimenten weer niet wilden lukken,… Kortom, bedankt dat je er altijd bent voor mij!

Contents

List of figures

List of tables

List of abbreviations

ACF autocorrelation function
AIDS acquired immunodeficiency syndrome
ALEX alternating laser excitation
APD avalanche photodiode
AOM acousto-optical modulator
AOTF acousto-optical tunable filter
BC back-complemented
CA HIV-1 capsid
CCD HIV-1 integrase catalytic core domain
CCF cross-correlation function
CD circular dichroism
(c)DNA (copy) deoxyribonucleic acid
CHO Chinese hamster ovary cells
CPM counts-per-molecule (kHz)
CTD HIV-1 integrase C-terminal domain
CW continuous wave
DIC differential interference contrast
dsDNA double stranded DNA
eBFP enhanced blue fluorescent protein
eCFP enhanced cyan fluorescent protein
eGFP enhanced green fluorescent protein
eYFP enhanced yellow fluorescent protein
FCS fluorescence correlation spectroscopy
FCCS fluorescence cross-correlation spectroscopy
FDA United States of America Food and Drug Administration
FLCS Fluorescence lifetime correlation spectroscopy
FLIM Fluorescence lifetime imaging microscopy
FRAP Fluorescence recovery after photobleaching
FRET Förster resonance energy transfer
GdHCl Guanidium hydrochloride
H1 Histone H1
HAART Highly active antiretroviral therapy
HDGF Hepatoma derived growth factor
HeLa Henrieta Lacks
HFT Haupt Farb Teiler, primary (dichroic) color splitter
HIV human immunodeficiency virus
HP1 Heterochromatin protein 1
HRP HDGF related protein
hRPB1 human RNA polymerase subunit B1
HSE heat shock element
IBD integrase binding domain
I(C)CS image cross-correlation spectroscopy
IN HIV-1 integrase
LEDGF/p75 Lens epithelium derived growth factor
LSM Laser scanning microscopy
LTR Long terminal repeat
MA HIV-1 matrix protein
MDF molecule detection function
miRNA microRNA
MLL mixed lineage leukemia
MLV Moloney Murine leukemia virus
MQ MilliQ water
mRNA messenger ribonucleic acid
MSD mean squared displacement
NA numerical aperture = $n.\sin\alpha$ with α the angular aperture
NC HIV-1 nucleocapsid
NFT Neben Farb Teiler, secondary (dichroic) color splitter
NLS nuclear localisation signal
NMR nuclear magnetic resonance
NTD HIV-1 integrase N-terminal domain
NUP98 nuclear pore complex protein 98
PBS phosphate buffered saline
PC4 positive cofactor 4
PCH photon counting histogram
PCR polymerase chain reaction
PDB protein databank
PIC HIV-1 pre-integration complex
PIE pulsed interleaved excitation

PogZ pogo transposable element derived protein with a Zinc finger
PWWP Pro-Trp-Trp-Pro motif containing domain
(m)RFP (monomeric) red fluorescent protein
RIC(C)S raster image (cross-)correlation spectroscopy
RNA ribonucleic acid
RNAPII RNA polymerase II
RT HIV-1 reverse transcriptase
s.d./SD standard deviation
SPAD single-photon avalanche diode
ssDNA single stranded DNA
STED stimulated emission depletion
STRE stress related element
SW-FCCS single wavelength FCCS
SWI/SNF switch/sucrose non-fermentable
TFFCS tunable focus FCS
TCSPC time-correlated single photon counting
TRN-SR2 transportin-SR2
TTTR time-tagged time-resolved

List of symbols

B	brightness
B_g	B of the green particle in green channel
B_r	B of the red particle in red channel
B_c	B of the green particle in red channel
D	diffusion coefficient ($\mu m^2/s$)
D_r	rotational diffusion coefficient
E	FRET efficiency
F	fraction
[G]	free concentration of protein G
$[G]_0$	total concentration of protein G
[GR]	concentration of the G-R complex
G(τ)	normalized fluorescence (fluctuation) autocorrelation function
$G_g(\tau)$	green ACF
$G_r(\tau)$	red ACF
G(0)	correlation function amplitude
$G_{CC}(\tau)$	normalized fluorescence (fluctuation) cross-correlation function
$G_D(\tau)$	diffusion term of the ACF
I(τ)	normalized fluorescence intensity autocorrelation function
K_{ass}	equilibrium constant for association
K_{diss}	equilibrium constant for dissociation
k_{off}	rate constant for dissociation
k_{on}	rate constant for association
k^*_{on}	$=k_{on}[sites]_0$; pseudo first-order rate constant for association
M	molar
	μM micromolar
	nM nanomolar
M_r	relative molecular mass
n	refractive index
N	average particle number
N_g	average 'only-green' particles
N_r	average 'only-red' particles
N_{gr}	average 'green-and-red' particles
$N_{g,total}$	$= N_g + N_{gr}$
$N_{r,total}$	$= N_r + N_{gr}$
N_{CC}	$= 1/G_{CC}(0)$
N_A	$= 6.022\times10^{23}$; Avogadro's number
Q	cross-talk parameter
[R]	free concentration of protein R
$[R]_0$	total concentration of protein R
r (Å)	intermolecular distance (Angström)
R	$= 8.314472\ JK^{-1}mol^{-1}$; gas constant
R_0 (nm)	Förster distance
r_H	hydrodynamic radius
s	second
	μs microsecond
	ms millisecond
	ns nanosecond
	h hours
S	$= \omega_2/\omega_1$; structure parameter
T (K)	temperature (Kelvin)
t_{bound}	$=1/k_{off}$; the bound time before dissociation
t_{free}	$=1/k_{on}$; the free time before association
V (fL)	volume (femtoliter, 10^{-15} L)
V_G	green volume
V_R	red volume
V_{CC}	cross-correlation volume
w/w	mass fraction, weight percent
α	anomaly parameter
Δ325	LEDGF/p75 lacking residues 1-325
η (cP)	viscosity (centipoises; mPa·s)
θ	rotational correlation time
λ (nm)	wavelength (nanometer)
ρ	density (kg/dm^3)
τ (ms)	correlation time (milliseconds)
τ_{diff}	diffusion time
τ_D	fluorescence lifetime of the donor
τ_{DA}	fluorescence lifetime of the donor in the presence of the acceptor
ω_1	radial radius of the confocal volume
ω_2	axial radius of the confocal volume

Samenvatting

Infectie met het humaan immuundeficiëntie virus (HIV), het virus dat het verworven immuundeficiëntie syndroom (AIDS) veroorzaakt, is ongeneeslijk; de zeer actieve antiretrovirale therapie is daarom een levenslange behandeling die gepaard gaat met een vermindering van de levenskwaliteit van de patiënt en met aanzienlijke ziektekosten. Het merendeel van geïnfecteerde mensen woont in ontwikkelingslanden en kan zich geen duur en doeltreffend therapeutisch regime veroorloven. Bovendien brengt antivirale resistentie alle doeltreffende behandelingen in gevaar. Basisonderzoek blijft dus nodig om de replicatiecyclus van HIV volledig te verstaan en om nieuwe doelwitten en nieuwe therapeutische strategieën te identificeren. LEDGF/p75, een menselijke transcriptionele co-activator, is belangrijk voor een efficiënte HIV-replicatie, door zijn directe en specifieke interactie met HIV-1 integrase, het virale enzym dat de insertie van het virale genoom in dat van de geïnfecteerde immuuncel katalyseert.

De grondreden voor dit doctoraal onderzoeksproject is het inzicht scheppen in de cellulaire functie van LEDGF/p75, zijn intracellulaire interactie met HIV-1 integrase en zijn rol voor de replicatie van het HIV-1. Om dit onderzoek uit te voeren hebben we een uitdagende interdisciplinaire aanpak gevolgd; kwantitatieve fluorescentietechnieken worden toegepast om het gedrag van eiwitten in levende menselijke cellen te bestuderen. Op die manier kunnen bepaalde moleculaire eigenschappen van LEDGF/p75 worden bepaald, die onmogelijk met normale biochemische onderzoekstechnieken kunnen worden onderzocht.

Fluorescentie is het fenomeen waarbij sommige stoffen licht kunnen uitzenden wanneer ze beschenen worden met licht van een lagere golflengte. Op die manier kan fluorescentie gebruikt worden als een 'markeerstift' voor biomoleculen. Confocale microscopische fluorescentietechnieken zijn het 'vergrootglas' dat we gebruiken om het gedrag van één soort biomolecule te onderzoeken in een levende cel die talloze andere biomoleculen bevat. Bijvoorbeeld, fluorescentiecorrelatiespectroscopie (FCS) is een biofysische techniek waarmee de concentratie en de dynamische eigenschappen van biomoleculen kan worden gemeten.

Fluorescentiecrosscorrelatiespectroscopie (FCCS) laat toe de interactie van specifieke eiwitten te bestuderen, en hun bindingsaffiniteit te bepalen.

Het experimentele gedeelte van dit project start met een reeks *in vitro* controleëxperimenten. De FCS-techniek, die origineel ontwikkeld werd om de concentratie en dynamica van eiwitten te bepalen in een waterige oplossing, wordt dan toegepast in een aantal verschillende experimentele omstandigheden. Het effect van de cellulaire brekingsindex, de temperatuur van de oplossing, de dikte van het dekglaasje, detector en fluorochroom gerelateerde artefacten op de experimentele FCS-metingen wordt nagegaan en een geschikte methode om een tweekleuren opstelling te kalibreren wordt ontwikkeld. In dit deel van de thesis zal worden nagegaan onder welke omstandigheden de techniek goed presteert wanneer kwantitatieve metingen moeten worden uitgevoerd in levende cellen.

Daarna wordt de transcriptionele co-activator LEDGF/p75 voor de eerste keer bestudeerd in real-time in zijn natuurlijke omgeving, in een levende menselijke celkern. Dit is mogelijk door LEDGF/p75 genetisch te merken met een groenfluorescent eiwit, eGFP, en de FCS-techniek toe te passen in cellen. We wijzen het dynamische gedrag van LEDGF/p75 toe aan interacties met het chromatine en geven een gedetailleerde mechanistische beschrijving van de kinetica van *in vivo* chromatinebinding. Wij zijn de eersten die zulke ultrasnelle kinetica van chromatinebinding rapporteren voor een transcriptiefactor. Het merendeel van de tijd is LEDGF/p75 op zoek naar mogelijke specifieke bindingsplaatsen op het chromatine, door middel van een dynamische genoomscanning. Tot slot geven we een betrouwbaar *in vivo* bewijs dat het PWWP-domein van LEDGF/p75, een gekend DNA/chromatine-bindend domein, een belangrijke bijdrage levert aan de associatie van LEDGF/p75 met chromatine.

In een volgend deel van de thesis werd de interactie tussen LEDGF/p75 en HIV-1 integrase bestudeerd in levende cellen door middel van de FCCS-techniek. Aangezien het intranucleaire complex van de twee eiwitten sterk aan het chromatine verankerd is, werden verschillende strategieën gevolgd om hun interactie toch te kwantificeren met FCCS: de chromatine-bindende domeinen van LEDGF/p75 werden verwijderd terwijl de interactie met IN werd verzekerd, het interactiecomplex werd verplaatst naar het cytoplasma van de cellen, of het belangrijkste chromatine-bindend domein werd op een specifieke manier gewijzigd. In de drie gevallen werd er een zodanige stijging in de mobiliteit van het complex waargenomen dat de interactie kon worden waargenomen met FCCS.

Heden worden deze interactietesten gebruikt om te screenen naar potentiële inhibitoren van deze interactie. Tot slot werden ook studies op de stoichiometrie van het interactiecomplex uitgevoerd en werd een eerste mechanistisch model opgebouwd, dat paradoxale meningen over transcriptionele co-activatie and targeting van HIV-1 door LEDGF/p75 verenigt.

Metingen op een eiwit door middel van fluorescentie heeft logischerwijs als vereiste dat het eiwit fluorescent gemerkt wordt. In het laatste deel van de thesis werd de nadruk gelegd op de fotofysica van de fluorescente merkers die in het onderzoek werden gebruikt, zijnde fluorescente eiwitten. De mate waarin de fotofysica van deze eiwitten de metingen van de eigenschappen van biomoleculen beïnvloedt werd onderzocht. Daartoe voerden we kwantitatieve FCCS-metingen uit op een translatiefusie van twee fluorescente eiwitten. Hierdoor kregen we meer inzicht in de biofysische mogelijkheden en beperkingen van de techniek, en we overdenken hoe de FCCS-techniek zou kunnen worden geoptimaliseerd in de toekomst. Tot slot hebben we de monomerische rood-fluorescerende eiwitten, die gebruikt werden voor de kwantitatieve FCCS-metingen, onderworpen aan een gedetailleerde spectroscopische *in vitro* analyse. We relateren onze observaties met kristallografische informatie en dragen bij aan de kennis over de vorming van donkere toestanden in deze eiwitten. Dit deel van het project zal bijdragen aan de toekomstige doelgerichte constructie van nieuwe stabiele rood fluorescerende eiwitten.

Summary

Infection with the human immunodeficiency virus (HIV), the virus that causes the Acquired immunodeficiency syndrome (AIDS), cannot be cured; the highly active antiretroviral therapy (HAART) is therefore lifelong and associated with a reduction in quality-of-life and considerable health care costs. Most infected people live in developing countries and cannot afford costly therapeutic regimens with effective antivirals. Moreover, emergence of antiviral resistance jeopardizes all effective treatment. Basic research thus remains necessary to fully understand the HIV replication cycle and to identify novel targets and novel therapeutic strategies. LEDGF/p75, a human transcriptional co-activator, is important for efficient HIV replication, through a direct and specific interaction with HIV-1 integrase (IN), the viral enzyme that catalyzes the insertion of the viral genome into that of the infected immune cell.

The rationale of this doctoral research project is to gain insight into the cellular function of LEDGF/p75, its intracellular interaction with HIV-1 integrase and its role for HIV-1 replication. A challenging interdisciplinary approach is followed; quantitative fluorescence techniques are applied to study the behaviour of proteins inside living human cells. This approach allows to determine molecular properties of LEDGF/p75 that are impossible to determine with standard biochemistry research tools.

Fluorescence is the phenomenon by which some substances emit light when irradiated with light of shorter wavelength. As such, it can be used as a 'magic marker' for biomolecules. Confocal microscopic fluorescence techniques are the 'magnifying glass' to visualize the intracellular behaviour of one type of biomolecule inside a living cell containing numerous other biomolecules. For example, fluorescence correlation spectroscopy (FCS) is a biophysical technique for measuring the concentration and the dynamic properties of biomolecules. Fluorescence cross-correlation spectroscopy (FCCS) allows to study the interaction of specific proteins and determine their binding affinity.

The experimental part of the project starts with a series of *in vitro* control experiments. The FCS technique, originally developed to measure protein concentrations and dynamics in aqueous environment, is applied in a range of different experimental conditions. The effects of the cellular refractive index, temperature of the solute, cover glass thickness, detector and fluorochrome related artefacts on the experimental measurements are verified and an adequate method for calibrating a two-color setup is developed. In this part of the thesis, we verify under which circumstances the technique performs well for quantitative measurements inside living cells.

Then, transcriptional co-activator LEDGF/p75 is studied for the first time in real-time in its natural condition, inside a living human cell nucleus. This is done by genetically labelling LEDGF/p75 with a green fluorescent protein, eGFP, and applying FCS in cells. We pinpoint the dynamic behaviour of the protein to chromatin binding and provide a detailed mechanistic description of the kinetics of *in vivo* chromatin binding. We are the first to report ultrafast kinetics of chromatin binding ($k^*_{on} > 1.9\times10^3$ s^{-1} and $k_{off} > 38.6$ s^{-1}) for a transcription factor. During the majority of its nuclear residence, LEDGF/p75 is thus dynamically scanning the genome, in search for possible specific binding sites. Finally, we provide solid *in vivo* evidence that the PWWP-domain of LEDGF/p75, a known DNA/chromatin binding motif, contributes strongly to chromatin association of LEDGF/p75.

In a next part of the project, we study the IN-LEDGF/p75 interaction in living cells with the FCCS technique. Since their intranuclear complex is strongly tethered to chromatin, different strategies are followed to probe the interaction with FCCS: the chromatin binding domains of LEDGF/p75 are removed while conserving the interaction with IN, the interaction complex is retargeted to the cytoplasm, or the major chromatin binding domain of LEDGF/p75 is specifically altered. In every case, the increase in mobility of the complex allowed to quantify the interaction with FCCS. Currently, this assay is used to screen for possible inhibitors of the interaction. Finally, we also performed interaction studies on the stoichiometry of the complex, and built a first mechanistic model that unifies paradoxical notions of transcriptional co-activation and HIV-1 targeting by LEDGF/p75.

Fluorescence measurements logically require the protein of interest to be fluorescently labelled. In the last part of the thesis, we focus on the photophysics of the fluorescent labels that were used throughout the project, i.e. fluorescent proteins, and try to assess to what

extent the photophysics influences measurements on actual biological molecules. We performed a quantitative FCCS analysis of a genetic fusion of two fluorescent proteins. We gained more insight in the biophysical possibilities and limitations of the method, and we speculate on how the method could be improved in the future. Finally, we subjected monomeric red fluorescent proteins (mRFPs), that are used for quantitative FCS and FCCS measurements, to a detailed spectroscopical analysis *in vitro*, related our results with structural information on RFPs and contributed to the knowledge on dark state formation. This part of the project will hopefully lead to the goal-oriented engineering of RFPs towards more stable fluorescence.

Part I – Rationale and objectives

The rationale of this doctoral research project is to gain insight into the intracellular function of human transcriptional co-activator LEDGF/p75, its interaction with HIV-1 integrase and its role for HIV-1 replication. To investigate this, we apply confocal spectroscopy techniques such as fluorescence correlation (FCS) and cross-correlation spectroscopy (FCCS), but also other quantitative fluorescence techniques inside living cells.

A first objective is to investigate the performance of the FCS technique, originally designed to measure protein concentrations and dynamics in aqueous environment, for quantitative measurements inside living cells. We will do this by exploring experimental conditions in a controlled *in vitro* environment. Possible methodological artefacts, that would be left unnoticed when measuring directly inside living cells, can in this way be identified, controlled for or even eliminated.

A second objective is to characterize the known chromatin binding properties of LEDGF/p75 for the first time in real-time inside living cells, with FCS. By labelling LEDGF/p75 with a green fluorescent protein, eGFP, we can monitor dynamics of the protein in the cell with FCS. The kinetics of chromatin binding will be characterized by customizing the experimental setup. Finally, the role of a major chromatin interacting domain for targeting HIV-1 integrase to chromatin will be investigated, both with FCS, as with a complementary technique for measuring protein dynamics in the slow time scale.

A third objective is to apply the FCCS technique for probing the interaction of HIV-1 integrase and LEDGF/p75 for the first time in living cells. Since their intranuclear complex is strongly tethered to chromatin, different strategies are followed to probe the interaction with FCCS: the chromatin binding domains of LEDGF/p75 will be removed, LEDGF/p75 will be retargeted to the cytoplasm, or the major chromatin binding domain of LEDGF/p75 will be altered. Finally, the FCCS technique will be used in a quantitative manner to determine and compare protein-protein interaction affinities.

A fourth objective is to evaluate the FCCS technique for measuring protein-protein interactions quantitatively in living cells. By combining FCCS measurements on a control system, a genetic fusion construct of two fluorescent proteins, with complementary measurements performed by another technique for quantifying protein-protein interactions, we gain more insight in the biophysical fundamentals of FCCS. Finally, the experimental setup and the fluorochromes used for FCCS are critically evaluated and limitations and possible optimizations are discussed.

Finally, a last objective is the characterization of monomeric red fluorescent protein photophysics with FCS, in combination with other spectroscopical methods. We provide structural-functional insight in this matter by linking our observations with crystallographic information.

Part II – Introduction

Chapter 1. HIV-1 integrase and human transcriptional co-activator LEDGF/p75

This thesis starts with the biological background of the doctoral research project. The human immunodeficiency virus is introduced and explained in detail with an emphasis on HIV-1 integrase and human LEDGF/p75, the two main proteins that were studied during the project.

Chapter 2. Fluorescence correlation and cross-correlation spectroscopy

The interdisciplinary character of the project will become clear in the second chapter, when the two principal techniques that were used, fluorescence correlation and cross-correlation spectroscopy, are introduced. This detailed introductory chapter gives an overview of the biophysical and mathematical basics of these techniques, and novel trends in this field are also compared.

Chapter 1. HIV-1 integrase and human transcriptional co-activator LEDGF/p75

1.1. The HIV/AIDS pandemic

In 1981 the CDC (U.S. Centers for Disease Control and Prevention) was alarmed by an unusually high prevalence of rare opportunistic infections such as *Pneumocystis carinii* pneumonia and Kaposi's sarcoma among young homosexual men. This condition was named Acquired Immunodeficiency Syndrome (AIDS). In 1983, the group of Luc Antoine Montagnier identified a previously unknown retrovirus to be the causative agent of AIDS (Barre-Sinoussi et al., 1983). Later the virus was named Human Immunodeficiency Virus type 1 (HIV-1). Between 1983 and 2007, 25 million people have died from the consequences of HIV-1 infection. In 2007 there were an estimated 33 million people in the world living with HIV-1, two-thirds of whom lived in Southern Africa (Figure 1.1). While the prevalence of HIV-1 has stabilized since 2000, the total number of people living with HIV-1 continues to increase.

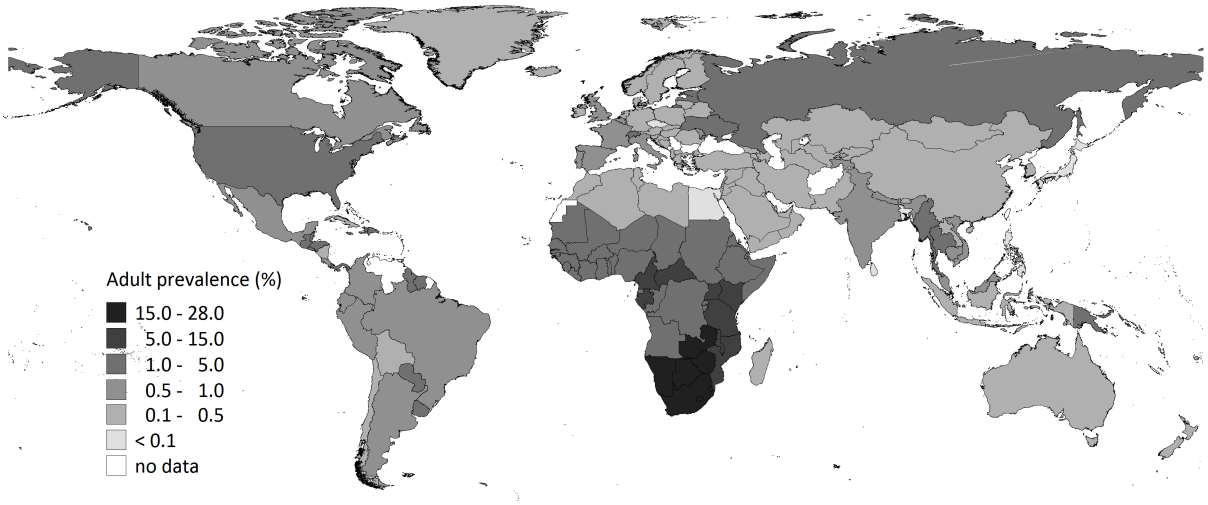

Figure 1.1 **Prevalence of HIV in the world** (reproduced from the 2008 UNAIDS report).

The HIV/AIDS pandemic has caused history's single sharpest reversal in human development. In the most heavily affected countries, it has reduced life expectancy considerably, increased poverty, shifted the size of populations, undermined national systems and weakened institutional structures. The epidemic cannot be reversed without reducing the rate of infections, therefore global HIV-1 prevention efforts, both on the economical, social and technical level, are essential. Recently though, a vaccine that lowered the chance of HIV infection with 26% has been reported, which holds promise for an effective prevention of HIV-1 in the future (Rerks-Ngarm et al., 2009).

HIV-1 is transmitted through the contact with bodily fluids. Heterosexual intercourse, injecting drug use, sex between men and mother-to-child transmission are the main driving forces of the epidemic. HIV-1 infection can be treated but still is incurable. The virus becomes irreversibly incorporated in the genetic material of the patient. A proof-of-principle study recently reported the possibility of excising HIV-1 from the genome (Sarkar et al., 2007), which might be a first step to a cure for HIV-1. Untreated, HIV infected patients usually progress to AIDS 9-10 years after the initial infection. Without treatment the survival time after progression to AIDS is less than one year. Clinical treatment of HIV-1 infection nowadays consists of the so-called highly active antiretroviral therapy (HAART). In this therapy patients are given a combination of different HIV targeting drugs in a strict schedule. HAART can extend the life expectancy of a patient by 20-50 years, but a large discrepancy with the life expectancy of HIV-1 negative individuals still exists. Moreover, side effects of the treatment can compromise the quality of life considerably. Continuous academic research on the biology and virology of HIV-1 is necessary, because it will aid the development of new

HIV-1 drugs or vaccines. Moreover, current treatment is complicated by drug toxicity and HIV-1 resistance development.

1.2. Virion and genomic structure of HIV-1

HIV-1 is a lentivirus of the family of retroviruses. The spherical ~145-nm virion (Briggs et al., 2003)(Figure 1.2) has an envelope consisting of a phopholipid bilayer with embedded complexes of a trimeric transmembrane protein (TM/gp41) tightly anchored to a trimeric surface protein (SU/gp120). The matrix protein (MA/p17) lines the inner part of the virion, providing an anchor for the gp120/gp41 complex and giving structural integrity to the virion.

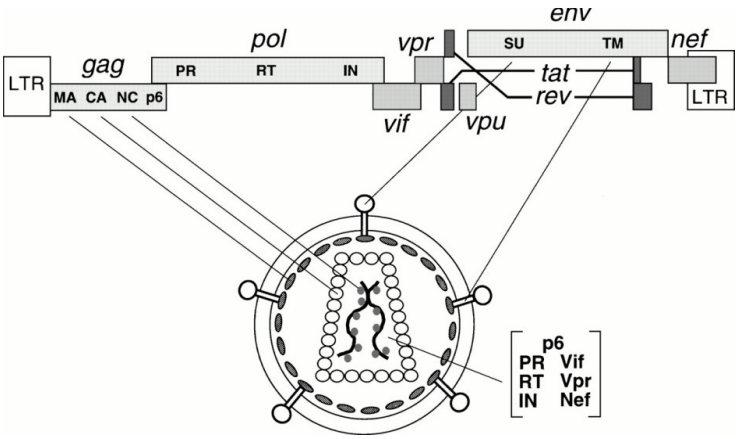

Figure 1.2 **Schematic structure of the HIV genome and virion** – the abbreviations are explained in the text (Frankel and Young, 1998).

A cone-shaped core consisting of 2000 copies of the capsid protein (CA/p24) is present in the virus. The virion contains the three viral enzymes, integrase (IN/p32), reverse transcriptase (RT/p51-p66) and protease (PR/p10) and the viral accessory proteins Vpr, Nef and Vif. Next to this, cellular proteins like BAF, Ini1 and cyclophilin A have been identified in the virion. The genome of HIV consists of two identical linear single stranded 9.2 kilobase (+)RNA strands, tightly associated with the nucleocapsid protein (NC/p6-p7) in the virion. Two copies of $tRNA^{Lys3}$, a primer for the reverse transcription of the genome, are hybridized with the genome in the virion. Structurally the genome consists of three large genes (*gag*, *pol* and *env*) and a number of smaller genes (*vif*, *vpr*, *vpu*, *tat*, *rev* and *nef*). The *gag* gene encodes for MA, CA and the NC proteins, the *pol* gene encodes for PR, RT and IN and the *env* gene encodes for the envelope proteins gp41 and gp120.

1.3. The HIV-1 replication cycle

HIV infects differentiated immune cells such as $CD4^+$ T cells and non-dividing macrophages and dendritic cells. The gp120 surface protein of HIV-1 specifically recognizes the CD4 receptor and a chemokine co-receptor on the immune cell. The type of chemokine receptor that is recognized by the virus, CCR5 or CXCR4, determines the tropism of the virus. After binding the gp41 protein undergoes a conformational change, triggering the fusion of the viral envelope with the host cell membrane and releasing the capsid into the cytoplasm (Figure 1.3). HIV can also enter the cell via receptor-mediated endocytosis and fusion of endosomes (Miyauchi et al., 2009). After uncoating of the capsid in the cytoplasm, reverse transcription of the RNA genome initiates. During this enzymatic reaction, the viral RNA is copied to double stranded copy DNA, flanked by two long terminal repeats (LTR's). This newly synthesized cDNA forms a complex with IN, RT, MA and Vpr and cellular proteins, a so-called pre-integration complex (PIC), which is transported actively into the nucleus by cellular transportin-SR2 (Christ et al., 2008) after being processed by IN in the first of two catalytic steps called 3'-processing. In the nucleus the viral genome is integrated into that of the host cell by IN in a second catalytic step referred to as strand transfer. Cellular repair mechanism finally repair the gaps flanking the integrated provirus.

The cellular transcription machinery transcribes the integrated provirus to mRNA starting at the promoter region in the 5'-LTR. Tat greatly enhances this process. In a next step the full-length RNA is either spliced and transported to the cytoplasm or directly transported to the cytoplasm, where new viral proteins are synthesized. This transport step is regulated by Rev. The Env transmembrane protein is synthesized in the endoplasmatic reticulum and directed to the cell membrane by vesicular transport. The Gag and Gag-Pol polyproteins are directly targeted to the cell membrane through the N-terminal, myristoylated MA protein. At the cell membrane, new viral particles are assembled. These new virions contain the polyproteins Gag and Gag-Pol, Vif, Vpr, Nef and the genomic RNA. Vpu promotes CD4 receptor downregulation, to prevent association with the viral receptor upon translation in the endoplasmatic reticulum. Nef promotes endocytosis and degradation of cell surface CD4 receptors. The exact function of Vif is less well known although it is important for the production of highly infectious mature virions. Vpr has a function in the nuclear uptake of the viral complex and induces cell cycle arrest in G2 phase. After budding from the cell, Gag

and Gag-Pol are proteolytically cleaved to MA, CA, NC, PR, RT and IN during maturation and the virion obtains its typical morphology and becomes infectious.

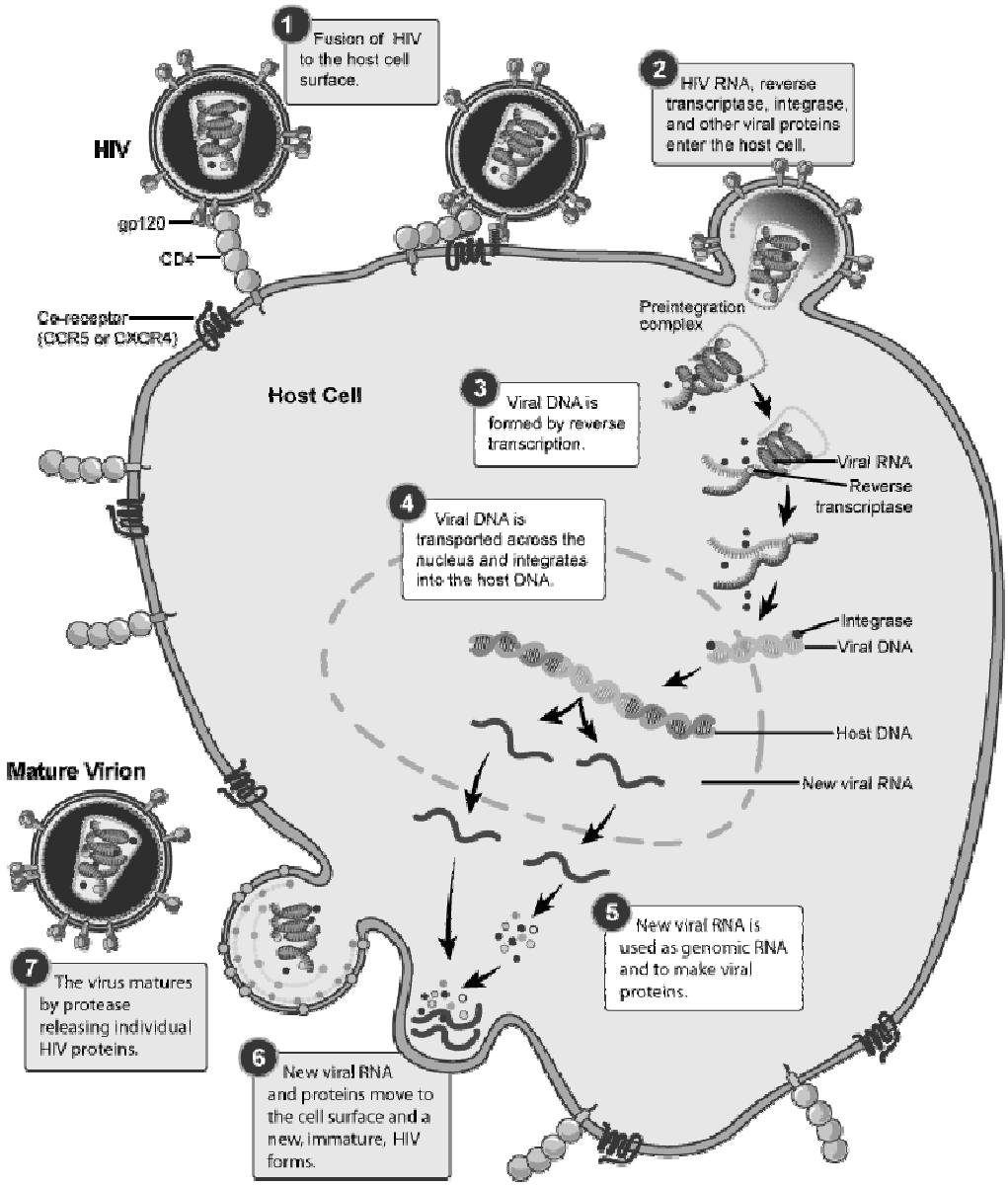

Figure 1.3 **The replication cycle of HIV** – (1) After binding to the host cell membrane, (2) the content of the virus is released in the cytplasm. (3) Reverse transcription converts the ssRNA viral genome into a cDNA copy. (4) After nuclear import, the viral genome is integrated into the host genome. (5) Transcription of the provirus allows for new viral genetic material and viral proteins to be synthesized. (6) New virions assemble at and bud from the cell membrane and (7) after maturation, infectious viruses are formed (credit: NIAID).

1.4. HIV-1 integrase

HIV-1 integrase (IN) is one of three essential retroviral enzymes present in the HIV-1 virion. IN (relative molecular mass 32 kDa) belongs to the nucleotidyl transferases (E.C.2.7.7.), enzymes that cut and paste DNA by direct transesterification. During HIV-1 replication, IN performs the enzymatic integration of the viral genome into the host genome. Mechanically and structurally, retroviral integrases resemble bacteriophage Mu and Tn5 transposases (Ljungquist et al., 1979; Reznikoff, 2003). During the transcription of new viral RNA, the mRNA of IN is part of the gag-pol mRNA. Translation is only possible when an occasional -1 frame shift occurs during translation of the gag gene. The viral enzyme protease finally releases IN by cleaving the Gag-pol polyprotein.

1.4.1. Enzymatic activity

IN catalyzes the integration of the HIV-1 genome in two enzymatic steps called the 3'-processing reaction and the strand transfer reaction (Figure 1.4). After reverse transcription in the cytoplasm, the double stranded viral cDNA in the pre-integration complex (PIC) is flanked by two identical Long Terminal Repeats (LTR's). IN now recognizes the two 5'-CAGT-3' consensus sequences at the cDNA termini and cleaves off the pGT dinucleotide, creating a recessed 3'-OH. This reaction is a nucleophilic substitution with either H_2O or the terminal 3'-OH as the attacking nucleophile. After nuclear import of the PIC, the two recessed 3'-OH groups perform a nucleophilic attack on the host chromosomal DNA, catalysed by IN. More specifically, the two complementary strands of the target DNA are linearized at a five base pair distance by the reaction and the integrated viral DNA becomes flanked by two single stranded 5-base gaps. These gaps are finally repaired by host DNA repair mechanisms after which a stably integrated provirus is formed.

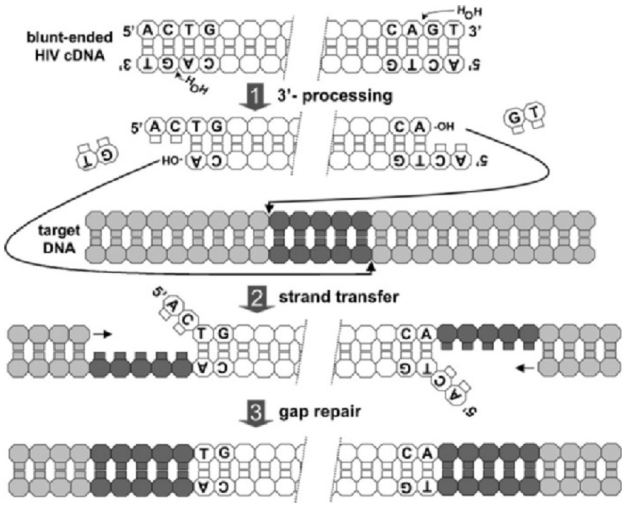

Figure 1.4 **Enzymatic steps catalysed by HIV-1 integrase during HIV replication** – In the first the 3'-processing step, IN binds to the consensus 5'-CAGT sequences at the LTR termini of the viral cDNA and cleaves off 5'-CA. During the second strand transfer step, IN inserts the two recessed LTR-ends at a 5-bp distance into the host genome. Host cell mechanisms repair the nicked ssDNA (Poeschla, 2008).

1.4.2. Structure

Like all retroviral integrases, IN contains three domains: an N-terminal domain (NTD), a catalytic core domain (CCD) and a C-terminal domain (CTD) (Figure 1.5). Crystallographic information is available for the NTD, CCD and CTD, for the fusion NTD-CCD and for the fusion CCD-CTD, but not for the full-length protein (Dyda et al., 1994; Lodi et al., 1995; Cai et al., 1997; Chen et al., 2000; Wang et al., 2001). The NTD (residues 1-52) forms a zinc-finger motif by complexation of His-12, His-16, Cys-40 and Cys-43 with a Zn^{2+} ion and forms a stable dimer. This motif is very well conserved among retroviral integrases and often makes up the DNA binding part of transcription activators. In the crystal of the NTD the helix normally used for DNA binding lies at the dimer interface (Cai et al., 1997) and is likely involved in forming the quaternary structure of IN. The CCD (residues 53-212) contains the triad of catalytic residues, Asp-64, Asp-116 and Glu-152, that bind a divalent metal ion such as Mg^{2+} or Mn^{2+} (Dyda et al., 1994; Jenkins et al., 1995) and that make up the catalytic site of the enzyme. The CCD is dimeric in solution. The CTD (residues 213-288) is least conserved among retroviral integrases. It shows homology with Src Homology 3 (SH3) domains, which play a role in protein-protein interactions in signal transduction pathways (Eijkelenboom et al., 1995). The CTD has been proposed to play a role in the oligomerization of IN (Lutzke and Plasterk, 1998).

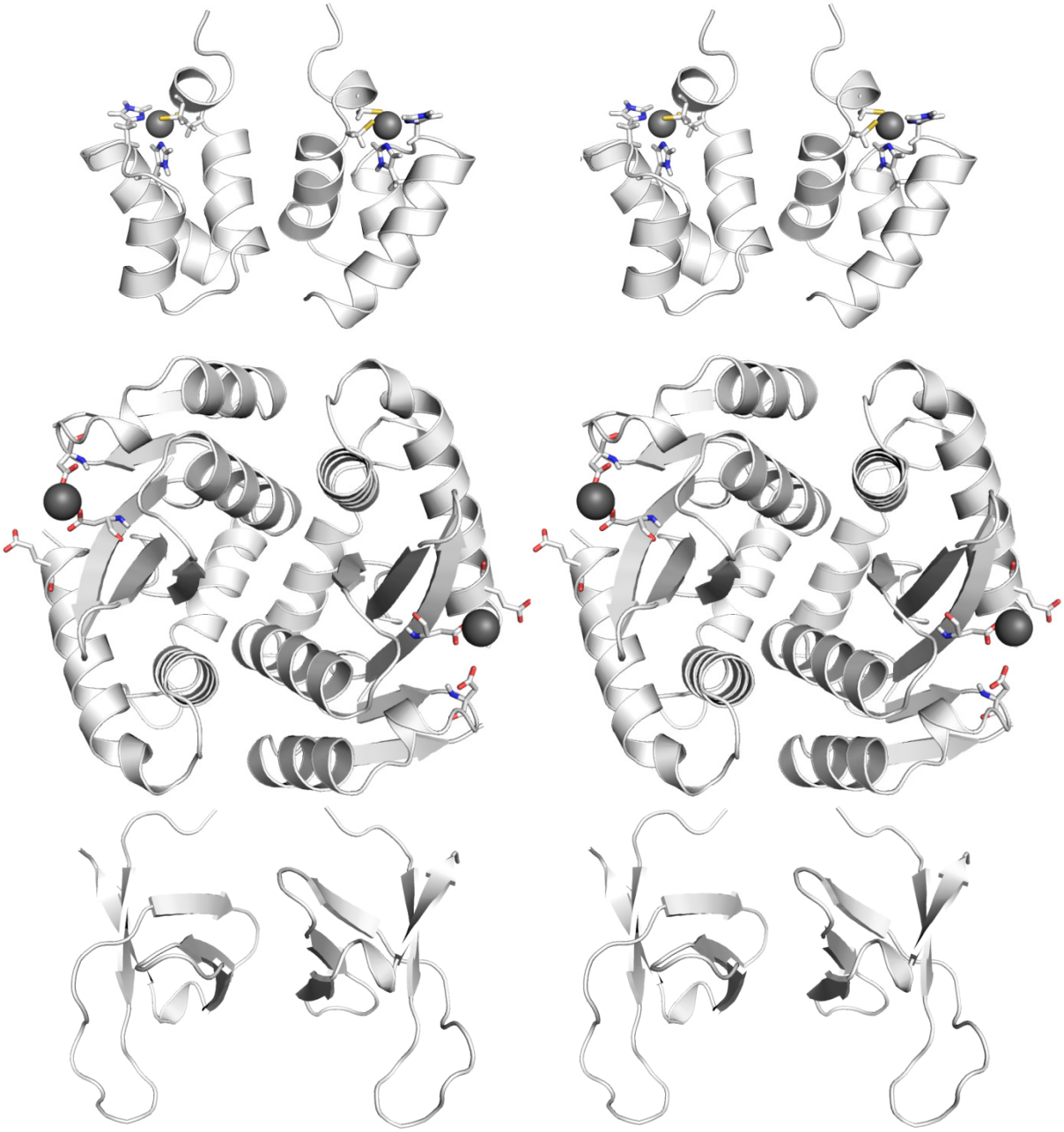

Figure 1.5 **Stereo representations of the crystallographic structures of the three domains of HIV-1 IN** (top) The dimeric N-terminal domain (residues 1-52), with a Zn^{2+} ion (in blue) bound to residues His-12, His-16, Cys-40 and Cys-43, shown as sticks (PDB id. 1WJC). (middle) The dimeric catalytic core domain (residues 53-212), with a Mg^{2+} ion (in green) bound to residues Asp-64, Asp-116 and Glu-152, shown as sticks (PDB id. 1BIU). (bottom) The dimeric C-terminal domain (residues 213-288) (PDB id. 1IHV)(produced with PyMOL Molecular Graphics System (2002, DeLano Scientific, Palo Alto, CA, USA)).

1.4.3. Stoichiometry

The individual domains of IN dimerize *in vitro*, as observed by crystallography. Moreover, the distance between two active sites in a IN-CCD dimer is larger than the five base pair spacing between the adjacent integration sites of the two ends of the viral cDNA (Dyda et al., 1994). This would imply that the active form of IN capable of concerted strand transfer contains a higher order stoichiometry or an equilibrium between forms with different stoichiometries (Craigie, 2001). Indeed, *in vitro* at μM concentration IN is present as a dimer (Jenkins et al., 1996). Recently, it has been claimed, based on FRET studies that the affinity of two monomers in a IN dimer is extremely strong, with an equilibrium dissociation constant of about 66 pM (Tsiang et al., 2009). In addition, IN can multimerize in solution (Jenkins et al., 1996; Delelis et al., 2008). This multimerization can be shifted to a dimer-tetramer equilibrium with an equilibrium dissociation constant of 22 μM by introducing the F185K-C280S solubility mutations. Furthermore, an oligonucleotide dsDNA substrate mimicking the LTR ends can either increase the solubility of IN by overtitrating with the DNA or increase IN multimerization by overtitrating the DNA with IN (Deprez et al., 2000; Vercammen et al., 2002). Furthermore it is known from cross-linking studies that *in vitro* full-site concerted integration of a 2-LTR substrate can only be performed by a tetrameric IN, while the dimeric form of IN can only perform half-site integration, the processing and integration of only one viral end (Faure et al., 2005; Guiot et al., 2006). The IN tetramer moreover does not show strand activity when the substrate contains only one viral LTR, implying a sequential binding of a monomer or dimer of IN to each cDNA end, subsequent 3' processing and finally tetramerization of IN and strand transfer (Figure 1.6).

Palindromic DNA sequences stabilize the IN tetramer due to their inherent symmetry (Delelis et al., 2007). Finally, molecular modeling attempts and a single particle reconstruction of electron microscopy data from DNA-IN complexes provide suggestions for the 3-dimensional structure of the complex (Podtelezhnikov et al., 2003; Karki et al., 2004; Ren et al., 2007), although the real quaternary structure of IN to date remains elusive. In summary, the smallest stoichiometry of free IN likely is a dimer but for strand transfer at least a tetrameric IN is necessary.

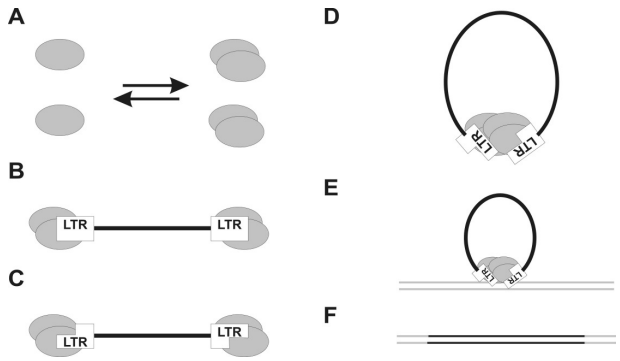

Figure 1.6 **IN sequential binding** (A) IN is present in a monomer-dimer equilibrium in solution. (B) the IN dimer binds to and (C) 3'-processes the viral cDNA. (D) Still bound to the cDNA IN tetramerizes and (E) integrates the viral genome into that of the host cell. (F) Finally, host cells mechanisms repair the 5' and 3' single stranded gaps in the integrated DNA (Faure et al., 2005).

1.5. HIV-1 integration and cellular components

Multiple steps in HIV-1 replication depend on human cell components. The CD4 immune receptor and some co-receptors are needed for entry, human tRNA primers for reverse transcription, the cytoskeleton for intracellular transport of the PIC (McDonald et al., 2002), TRN-SR2 for nuclear import of the PIC (Christ et al., 2008) and the cellular transcription machinery for proviral gene expression, to give a few examples. Specifically for HIV-1 integration, a number of human proteins have been proposed to play a role. Barrier-to-autointegration factor (BAF), a highly conserved protein with a role in nuclear structure organization and assembly, associates with PICs in cells, protects retroviral DNA against autointegration and has been shown to have a role in HIV-1 integration (Lee and Craigie, 1998; Lin and Engelman, 2003; Violot et al., 2003; Margalit et al., 2005). High Mobility Group Chromosomal protein A1 (HMGA1), a protein involved in transcription regulation by chromatin remodeling, can restore *in vitro* integration activity of salt stripped PICs (Farnet and Bushman, 1997; Reeves, 2001). Integrase interactor 1 (INI1/hSNF5), a component of the ATP-dependent chromatin remodeling mammalian SWI/SNF complex, interacts directly with IN, stimulates IN *in vitro* and is incorporated in HIV-1 virions (Kalpana et al., 1994; Muchardt and Yaniv, 1999; Yung et al., 2004). Moreover, overexpression of the IN binding domain of INI1 inhibits HIV-1 replication (Yung et al., 2001). Finally, Lens epithelium derived growth factor (LEDGF/p75), a human transcriptional co-activator (Ge et al., 1998a), was identified as an IN co-factor in a co-immunoprecipitation experiment (Cherepanov et al., 2003). We will discuss LEDGF/p75 in more detail in the following section.

1.6. The transcriptional co-activator LEDGF/p75

1.6.1. Discovery of LEDGF/p75

Lens epithelium derived growth factor (LEDGF) was originally identified as a 75-kDa growth and survival factor in human lens epithelial cells (Singh et al., 1998). Concurrently it was identified as a 75-kDa protein (p75) during purification of the PC4 transcriptional co-activator and was shown to interact with the VP16 activation domain and with components of the general transcription machinery (Ge et al., 1998a). Its alternative splice variant, p52, is a 333 amino acid protein that shares the first 325 amino acids with LEDGF/p75. LEDGF/p75 and p52 localize in different compartments in the nucleus (Nishizawa et al., 2001). Next to a broader range of interactions with transcription activators, p52 also functionally interacts with ASF/SF2 (alternative splicing factor/splicing factor 2), also known as SFRS1 (splicing factor, arginine/serine-rich 1) (Ge et al., 1998a; Ge et al., 1998b). LEDGF/p75 is a 530 amino acid protein derived from the *ledgf* gene, which is also known as the *psip1* gene, for PC4- and SF2-interacting protein 1 (Figure 1.7). Finally, LEDGF was also referred to as dense fine speckles 70-kDa (DFS70) protein in autoimmune diseases (Ochs et al., 2000). Since LEDGF is an abundant protein that stimulates survival in a wide range of cell types, its name is not well chosen (Singh et al., 2000b). Nonetheless, the currently accepted and unique acronym is LEDGF/p75.

LEDGF/p75 also has a function in oncogenesis. Through a direct interaction with the MLL/menin complex, LEDGF/p75 has been shown to be indispensable for MLL (mixed lineage leukemia) histone methyltransferase (HMT) dependent transcription and for MLL/menin dependent leukemic transformation (Yokoyama and Cleary, 2008). Furthermore, LEDGF/p75 is targeted by chromosomal translocations in leukemia that result in fusion with the nucleoporin NUP98 (Ahuja et al., 2000; Hussey et al., 2001; Grand et al., 2005; Morerio et al., 2005). Autoantibodies against LEDGF/p75 are present in certain autoimmune diseases and prostate cancer (Ganapathy et al., 2003; Daniels et al., 2005). During apoptosis LEDGF/p75 undergoes caspase-3 and -7 dependent cleavage, which abrogates the pro-survival function and enhances its immunogenicity (Wu et al., 2002). The function of LEDGF/p75 during HIV-1 replication will be discussed later.

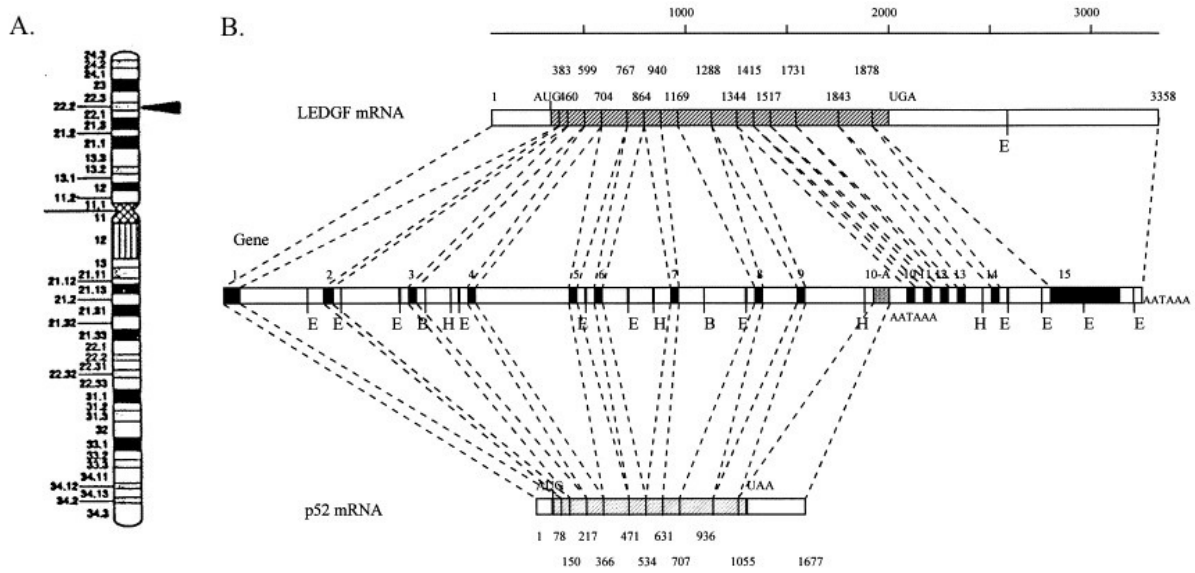

Figure 1.7 **The location and organization of the *ledgf* gene** – The *ledgf* gene is located on chromosome 9 area 9p22.2, contains 15 exons and 14 introns and encodes for two proteins, LEDGF/p75 and its alternative splice variant p52 (Ge et al., 1998a; Singh et al., 2000a) A. Organization of chromosome 9. Black arrow = position of the ledgf gene. B. Organization of the ledgf gene and the LEDGF/p75 and p52 mRNA. Ledgf gene: exons and introns are represented by resp. black and white colors, the exons are numbered. E, B and H indicate restriction sites for resp. EcoRI, BamHI and HindIII. Polyadenylation signals are indicated with AATAAA. LEDGF/p75 and p52 mRNA: coding and non-coding parts are represented by resp. gray and white colors. AUG: start codon; UAA: stop codon (Singh et al., 2000a).

1.6.2. Chromatin binding domains in LEDGF/p75

LEDGF/p75 belongs to the family of Hepatoma derived growth factor (HDGF) related proteins (HRP's), which includes p52, HRP-1, -2, -3 and -4 (Izumoto et al., 1997; Ikegame et al., 1999; Dietz et al., 2002). Hepatoma derived growth factor (HDGF) is a protein with mitogenic properties originally identified in human hepatoma derived cell lines but ubiquitously expressed in several other cell lines (Nakamura et al., 1994). Besides the presence of a well-conserved N-terminal PWWP-domain, HRPs show very little homology (Izumoto et al., 1997; Stec et al., 2000; Singh et al., 2006). Characteristic of PWWP domains is a conserved although not invariant Pro-Trp-Trp-Pro motif (Stec et al., 2000). PWWP domains are generally involved in protein-protein interactions regulating the chromatin structure (Stec et al., 2000). Although structural information on the PWWP domain of LEDGF/p75 is not known, a comparison of the known structures of the PWWP domains of HDGF, Dnmt3b and mouse HRP-3 showed a high degree of similarity (Qiu et al., 2002b; Sue et al., 2004; Nameki et al., 2005; Lukasik et al., 2006). In addition, it has been shown that the PWWP

domain of HDGF can form domain swapped dimers *in vitro* (Sue et al., 2007) and that SUMO (small ubiquitin like modification) of HDGF occurs in cells (Thakar et al., 2008). In LEDGF/p75 a tripartite element consisting of the two AT-hooks and the NLS cooperates with the PWWP-domain when binding to DNA/chromatin, as has been shown *in vitro* (Llano et al., 2006b) and in vivo (Turlure et al., 2006) (Figure 1.8). In addition, eGFP-PWWP can interact with mitotic chromosomes and deletion of only this domain from full-length LEDGF/p75 weakens its affinity for chromatin *in vitro* (Llano et al., 2006b) and in vivo (Turlure et al., 2006; Llano et al., 2006b), underlining its contribution to chromatin-binding. Finally, PWWP domain residues important for binding to an *in vitro* chromatinized template were recently identified (Shun et al., 2008).

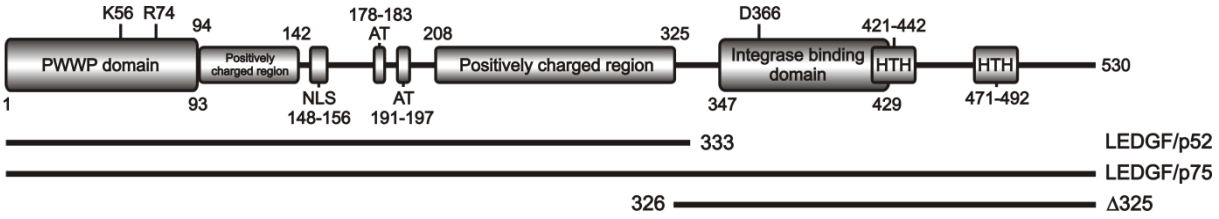

Figure 1.8 **Schematic representation of LEDGF/p75** – LEDGF/p75, its alternative splice variant LEDGF/p52 (Singh et al., 2000a) and a deletion mutant of LEDGF/p75 lacking residues 1-325 (Δ325)(De Rijck et al., 2006). NLS = nuclear localisation signal, AT = AT-hook domains, HTH = predicted Helix-Turn-Helix motifs (Llano et al., 2006b).

Next, LEDGF/p75 directly interacts with heat shock elements (HSE) and stress related regulatory elements (STRE) to promote expression of stress-related genes, such as Hsp27, αB-crystallin, alcohol dehydrogenase (ADH), anti-oxydant protein 2 (AOP2) (Fatma et al., 2001; Singh et al., 2001; Shinohara et al., 2002; Sharma et al., 2003; Fatma et al., 2004). It activates stress-response genes by virtue of the N-terminal PWWP domain binding to a STRE in the promotor of this gene. Furthermore it can activate stress-response genes by binding of its C-terminus (418-530), containing two predicted helix-turn-helix domains, to the HSE's of the same promotor (Singh et al., 2001; Singh et al., 2006). Finally, LEDGF/p75 contains a large percentage of charged residues (39.4%) and contains two defined regions with positively charged residues, that are involved in electrostatic interactions with DNA/chromatin (Llano et al., 2006b).

1.6.3. A protein-protein interaction domain in LEDGF/p75

Next to the PWWP domain, LEDGF/p75 contains another conserved structural domain that was originally identified as the IN binding domain (IBD) (Cherepanov et al., 2004). Indeed, this domain of LEDGF/p75 strongly and specifically interacts with IN (Cherepanov et al., 2005a). NMR studies provided structural details on the IBD; it consists of 5 alpha helices that form a superhelical domain. The IBD is structurally homologous to HEAT (Huntingin, elongation factor 3, subunit A of phosphatase 2A, yeast PI3 kinase TOR) or ARM (*Drosophila melanogaster* protein Armadillo) repeats of various proteins, which usually are responsible for mediating protein-protein interactions (Cherepanov et al., 2005b). A co-crystal of IBD and IN-CDD showed that the IBD of LEDGF/p75 interacts specifically with IN by binding at the interface of the IN-CCD dimer (Cherepanov et al., 2005a) (Figure 1.9). Other human proteins have been identified that bind to the IBD. Firstly, JPO2, also known as RAM2, represses transcription of the gene encoding monoamine oxidase (MAO) by binding to SP1 sites in the core promoter of MAO (Chen et al., 2005a) and was independently identified in a yeast two hybrid screen for interactors of the Myc transcription factor (Huang et al., 2005). JPO2 interacts with LEDGF/p75 via the IBD and can compete with IN (Maertens et al., 2006). LEDGF/p75 colocalizes with JPO2 in interphase cells, tethers JPO2 to mitotic chromosomes and stabilizes intracellular levels of JPO2 (Maertens et al., 2006). The JPO2/IBD interaction surface is different from the IN-CCD/IBD interface (Bartholomeeusen et al., 2007). Secondly, the menin tumor suppressor, product of the *men1* gene and implicated in cancer and transcriptional regulation as a component of the MLL-HMT complex (section 1.6.1) interacts specifically with LEDGF/p75. The IBD and adjacent HTH-motif are necessary and sufficient to target LEDGF/p75 to the MLL/menin complex (Yokoyama and Cleary, 2008). Thirdly, PogZ (pogo transposable element derived protein with a Zinc finger), also interacts specifically with LEDGF/p75 through the IBD. PogZ is a domesticated transposase and carries a DDE-domain similar to the IN-CCD (Bartholomeeusen et al., 2009). PogZ also competes with IN for binding to LEDGF/p75, although the affinity is likely weaker (Bartholomeeusen et al., 2009). Finally, Cdc7 is a Ser/Thr kinase essential for the initiation of DNA replication throughout the S-phase. Its activity is controlled via interaction with a regulatory subunit, activator of S-phase kinase (ASK). LEDGF/p75 interacts with heterodimeric Cdc7/ASK via the IBD and stimulates Cdc7 kinase activity (Hughes et al., 2010). In summary, the IBD is a well conserved motif in

LEDGF/p75 that seems to function as an adaptor for interacting with multiple cellular target proteins.

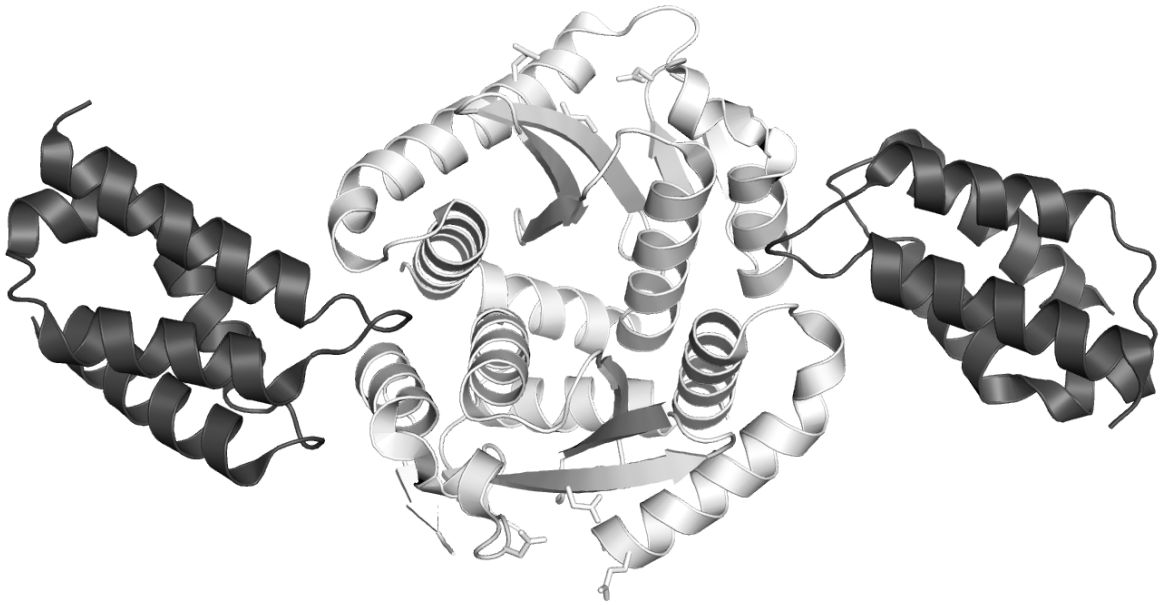

Figure 1.9 **The crystal structure of the dimeric catalytic core domain of IN (white) complexed to the IN-binding domain of LEDGF/p75 (gray)** – Residues important for the enzymatic activity of IN are shown as sticks (PDB id. 2B4J)(Cherepanov et al., 2005a)(produced with PyMOL).

1.7. LEDGF/p75 and HIV-1 replication

LEDGF/p75 is important for HIV replication. Potent knock-down or knock-out of LEDGF/p75 interferes with efficient HIV replication (Vandekerckhove et al., 2006; Zielske and Stevenson, 2006; Llano et al., 2006a; Marshall et al., 2007). Virus clones carrying mutations at the integrase-LEDGF/p75 interface are defective for HIV replication (Emiliani et al., 2005; Busschots et al., 2007). Overexpression of C-terminal fragments of LEDGF/p75 drastically inhibits HIV replication (De Rijck J. et al., 2006a). Resistant HIV strains selected in cells overexpressing those fragments, carry mutations which lie at the integrase-LEDGF interface but remain dependent on the co-factor (Hombrouck et al., 2007).

1.7.1. A tethering function of LEDGF/p75

Upon identification of LEDGF/p75 by co-immunoprecipitation of HIV-1 IN, it was shown to enhance the strand transfer activity of IN *in vitro* (Cherepanov et al., 2003). This effect is lentivirus specific (Busschots et al., 2005; Cherepanov, 2007). In section 1.6.2 we already discussed the literature on the chromatin binding of LEDGF/p75. In addition, LEDGF/p75 has been shown to greatly enhance the DNA binding of IN *in vitro* (Busschots et al., 2005). In

the living cell, IN and LEDGF/p75 co-localise in the nucleus and remain stably bound to chromosomes during mitosis (Maertens et al., 2003). Mutation of the NLS of LEDGF/p75 disrupts the nuclear localization of IN and RNAi-mediated knock-down of LEDGF/p75 results in a disappearance of the preferential nuclear localization of IN (Maertens et al., 2004; Llano et al., 2004b; Vandekerckhove et al., 2006). Based on all these experiments, a tethering function of LEDGF/p75 for the PIC was put forward. The N-terminal p52-part mediates DNA/chromatin-binding and the C-terminal Δ325-part can bind to IN through the conserved IBD, as such forming a molecular bridge between IN and the chromatin (Maertens et al., 2003; De Rijck J. et al., 2006b; Hombrouck et al., 2007). The apparent effect of LEDGF/p75 on the nuclear import of the pre-integration complex is most likely an indirect consequence of tethering (Maertens et al., 2004). Independent from the appreciable amount of research that has been performed on LEDGF/p75, to date no information is available on the strength of chromatin binding/tethering of LEDGF/p75 and if strong tethering of IN is a prerequisite of efficient integration. This is merely due to the fact that common *in vitro* techniques cannot assess interactions with chromatin quantitatively inside the cell.

1.7.2. A targeting function of LEDGF/p75

When LEDGF/p75 tethers IN to the chromatin, an entropic advantage for strand transfer is observed *in vitro*. In a cellular context however, integration is moreover targeted to actively transcribed regions in the genome. IN is the principle viral component responsible for this (Lewinski et al., 2006). LEDGF/p75 interacts with different components of the cellular transcription machinery and has even been shown to bind to specific sequences in promoter regions of certain genes (sections 1.6.1 and 1.6.2). Upon knockdown of LEDGF/p75 using RNAi, the integration pattern of HIV is perturbed (Ciuffi et al., 2005; Marshall et al., 2007; Shun et al., 2007). Furthermore, LEDGF/p75 responsive genes were identified by transcriptional profiling and found to be favored integration targets for HIV (Berry et al., 2006). LEDGF/p75 is thus an important human factor that targets HIV replication. In the absence of LEDGF/p75, however, through knockdown or knockout, integration is still biased towards actively described regions (Ciuffi et al., 2005; Marshall et al., 2007; Shun et al., 2007). Merely the more efficient integration in transcriptionally active and decondensed euchromatin in the nucleus could partially explain this; alternatively, another human or viral targeting factor for HIV-1 could exist. Hrp-2, for example, has been shown to be able to

interact specifically with HIV-1 integrase (Cherepanov et al., 2004). Its possible function in targeting HIV in the absence of LEDGF/p75 remains to be shown.

1.7.3. A coordinating function of LEDGF/p75

LEDGF/p75 was originally discovered during co-immunoprecipitation of HIV-1 IN. It bound with two copies to a tetramer of IN (Cherepanov et al., 2003). Later work showed that LEDGF/p75, and the IBD, could indeed promote IN tetramerization, which positively influences the strand transfer activity of IN (McKee et al., 2008). LEDGF/p75 IBD derived peptides that had an effect on the oligomerization of IN were shown to inhibit IN enzymatic activity (Hayouka et al., 2007). Structurally, the IBD of LEDGF/p75 binds at the dimer interface of an IN-CCD dimer (Cherepanov et al., 2005a). Importantly, contacts between the IN-NTD and the IBD have been suggested from a co-crystal of HIV-2 IN-NTD-CCD and IBD (Hare et al., 2009b). Since the function of the NTD has been shown to be promoting tetramerization and increasing enzymatic activity of IN, the effect of the IBD on the IN stoichiometry could be related to these contacts (Zheng et al., 1996; Lee et al., 1997). Cryo-EM studies provided a first *in vitro* model of a tetrameric IN complexed with two LEDGF/p75 molecules (Michel et al., 2009). Finally, a crystal structure of maedi-visna virus IN NTD-CCD in complex with the IBD finally suggested that a catalytically active IN tetramer can be formed by domain swapping of the NTD's of two monomers that are closest to each other in a IN tetramer (Figure 1.10)(Hare et al., 2009a). Important in this crystal is the presence of 4 unidentical binding sites for the IBD, two with strong affinity at the tetramer interface that make contacts both with the NTD as with the CCD, and two on the outside of the tetramer, of weaker affinity.

In conclusion, although a crystal structure of full-length HIV-1 IN in complex with LEDGF/p75 has not yet been resolved, the CCD of IN likely induces a dimer of IN, and these dimers then constitute a tetramer by domain swapping of NTDs, interaction which are stabilized by LEDGF/p75.

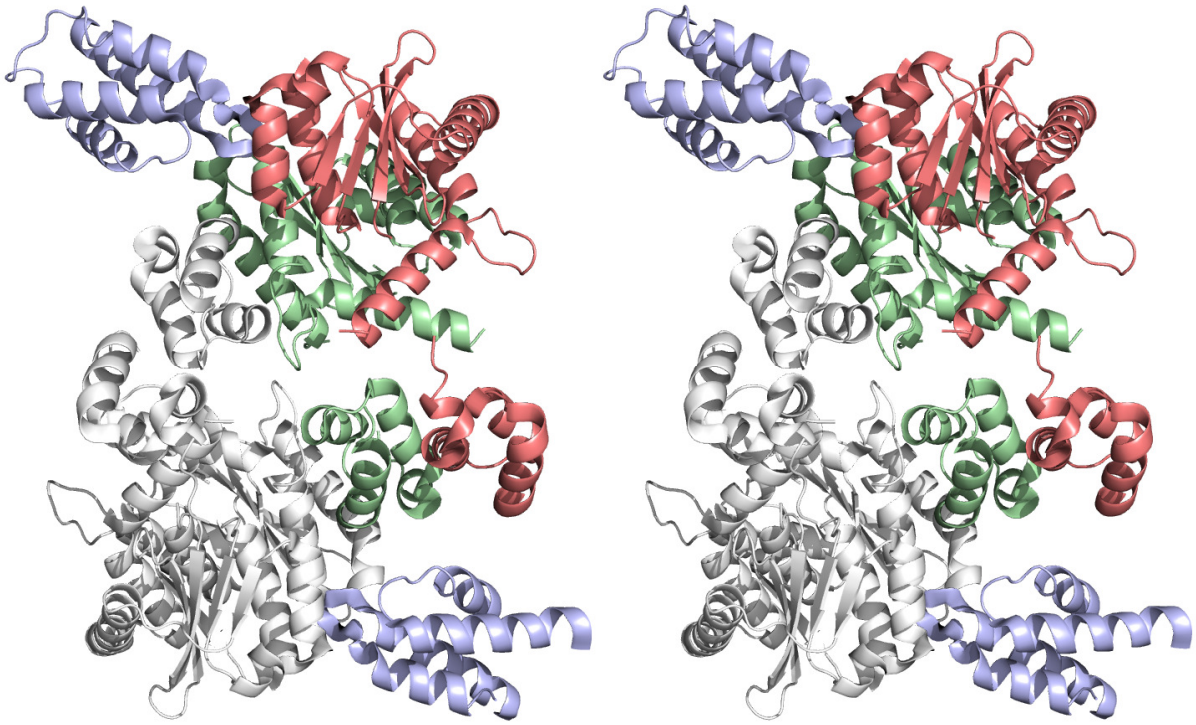

Figure 1.10 **Co-crystal structure of maedi-visna IN (NTD-CCD) and LEDGF/p75 (IBD) in stereo view** – A tetrameric IN is shown, of which one dimer is shown in red/green and one dimer in gray. Two of four IBDs are shown in blue. The small domain swapped colored NTDs of the top dimer make extensive contacts with the lower dimer. Whereas the bottom IBD is bound to both the IN-CCD dimer as the IN-NTD from another IN dimer, the top IBD is only bound to the IN-CDD dimer (Hare et al., 2009a) (produced with PyMOL).

1.7.4. A protective function of LEDGF/p75

Finally, IN expression in cells is stabilized by the presence of LEDGF/p75. LEDGF/p75 was shown to protect IN against proteasomal degradation (Llano et al., 2004a) and has recently also been shown to have a protective function for other proteins (Maertens et al., 2006). This protection seems independent from nuclear relocalisation or chromatin tethering. Likely, the higher order stoichiometry of the LEDGF/p75-IN complex prevents IN from being targeted for degradation in cells.

1.8. Novel ways to inhibit HIV-1 replication

Treatment of HIV infection is possible with highly active anti-retroviral therapy (HAART). Patients following HAART treatment take 'cocktails' of antiretroviral drugs that target different steps in the HIV replication cycle; viral fusion, reverse transcription, integration and proteolytic cleavage of newly synthesized viral polyproteins. However, due to the high error

rate of the viral reverse transcriptase enzyme, on average about 1 error per 2000 synthesized bases, and the high turnover of viral particles, 10^9 to 10^{10} new particles are produced each day in a patient, the genotype of the virus can change remarkably easy. This genotypic change allows for the virus to become resistant against the therapy, if viral proteins targeted by the drugs are modified by the genotypic change. Currently, FDA approved HIV drugs can be divided into six classes, the nucleoside analog, nucleotide analog and non-nucleoside reverse transcriptase inhibitors (resp. NRTIs, NtRTIs and NNRTIs), protease inhibitors (PIs), entry or fusion inhibitors and the most recent class, the integrase inhibitors. LEDGF/p75 is important for efficient HIV replication and has been put forward as a novel target for HIV treatment.

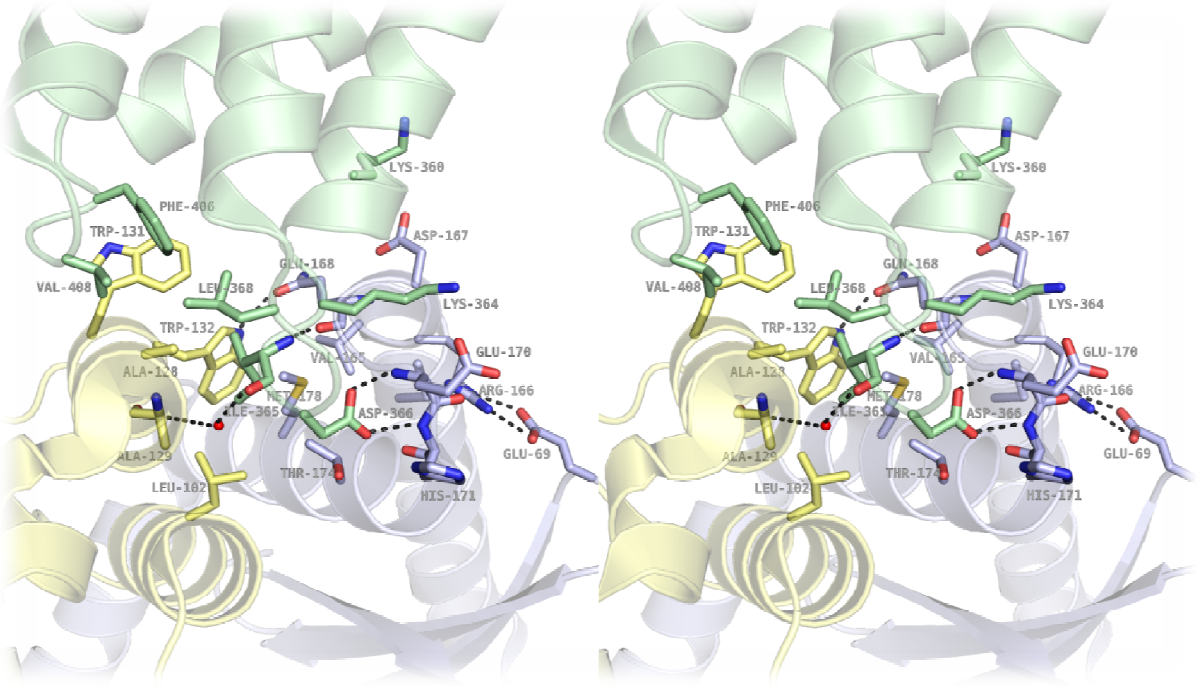

Figure 1.11 **Stereo zoom on the contact surface of a co-crystal of the IN-CCD dimer and LEDGF/p75-IBD** – Monomers in the IN-CCD dimer are shown in yellow and blue, the IBD is shown in green. Important residues are shown as sticks. Important H-bonds are indicated (produced with PyMOL).

Generally, drugs target small sites on a protein, such as an active site in an enzyme. Protein-protein interactions, such as the LEDGF/p75-IN interaction, generally have large contact surfaces, making it unattractable for the development of small-molecule inhibitors. In the particular case of the LEDGF/p75-IN interaction, however, the interhelical loop of the IBD protrudes into the IN-CCD dimer interface (Figure 1.11), creating a rather small contact

surface that could potentially be blocked by a small-molecule inhibitor. For example, it has been shown that single D366A, I365A or F406A mutations in IBD completely knock out the interaction with IN (Cherepanov et al., 2005b). Importantly, since LEDGF/p75 is not subjected to the genetic variability of the virus, inhibitors can potentially be designed that are less jeopardized by resistance development. Drug design for the inhibition of this interaction is part of ongoing research. Finally, the chromatin binding of LEDGF/p75 and its possible role in tethering the pre-integration complex to chromatin has been put forward as an important function for HIV-1 replication (De Rijck et al., 2006). Inhibition of chromatin binding has been shown to have profound effects on HIV-1 replication and could possibly constitute another important step at which HIV replication could be targeted.

Chapter 2. Fluorescence correlation and cross-correlation spectroscopy

2.1. Confocal microscopy

Optical microscopy is used to produce a magnified image of a microscopic sample using light. In a conventional epi-fluorescence microscope the excitation light from a laser or lamp is focused at the back aperture of the objective. This excitation light enters the sample defocused, illuminating the complete field-of-view homogenously, hence the name 'wide-field' illumination (Figure 2.1 left).

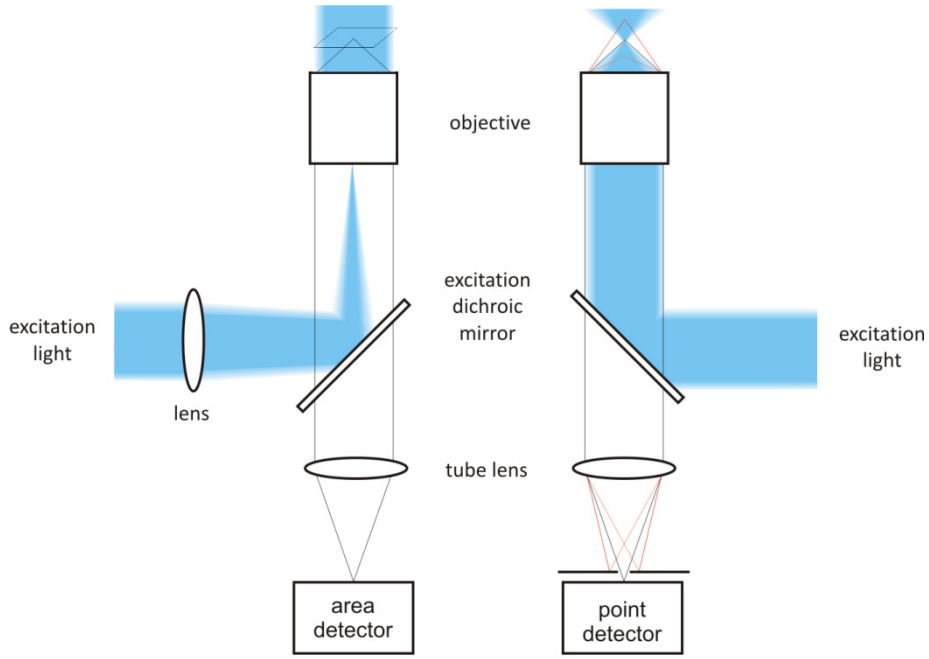

Figure 2.1 **The wide-field versus the confocal principle** – In blue is the excitation light. The thin black or colored lines represent the optical path of a sample in (black), above (red) or below (orange) the focal point of the objective. Under confocal excitation, the objective and pinhole allow for optical confinement of the observation spot..

Fluorescence light from the substrate passes through the objective again and is spectrally and spatially separated from the excitation light by a dichroic mirror placed in the optical path. An emission filter further selects the emission band of the fluorochrome, after which the emission is registered by a detector. This detector, which can be the human eye or a camera, registers an exact image of the substrate when it is positioned in an image plane that is conjugate to the objective front focal plane. Optical resolution in wide-field microscopy, defined as the smallest distance between two point sources in the sample that can still be distinguished, is equal to $d=0.6\lambda/NA$, with λ the wavelength and NA the numerical aperture of the objective lens. For a high-NA objective, the resolution of the registered image will thus roughly be half the wavelength used to observe the substrate.

In a confocal fluorescence microscope, a laser source provides stable monochromatic excitation light (Figure 2.1 right). This collimated light beam completely fills the back aperture of an objective with high numerical aperture. This creates an excitation spot spatially limited only by the wavelength of the light, a so-called diffraction-limited spot. Owing to the higher contrast, the optical resolution achieved on a confocal microscope is slightly better than in a wide-field microscope ($d=0.4\lambda/NA$). Fluorescence seen by the objective is separated from the excitation by a dichroic mirror and directed to a detector with high quantum yield after passing through an emission filter. Optical sectioning in the axial direction is achieved by placing a pinhole in a conjugated image plane of the optical path. Under widefield illumination this pinhole would simply act as a field diaphragm, limiting the image of the sample in X-Y to a spot not much bigger than the optical resolution (Figure 2.2 left). Under confocal excitation on the other hand only a diffraction-limited spot in the sample is illuminated (Figure 2.2 right). The pinhole will now act as a strong off-plane detection cut-off, since only light waves emanating from exactly the focal plane of the objective will be focused in the pinhole and thus pass through. The projection of the sample in confocal imaging mode is thus limited in both X and Y as in Z.

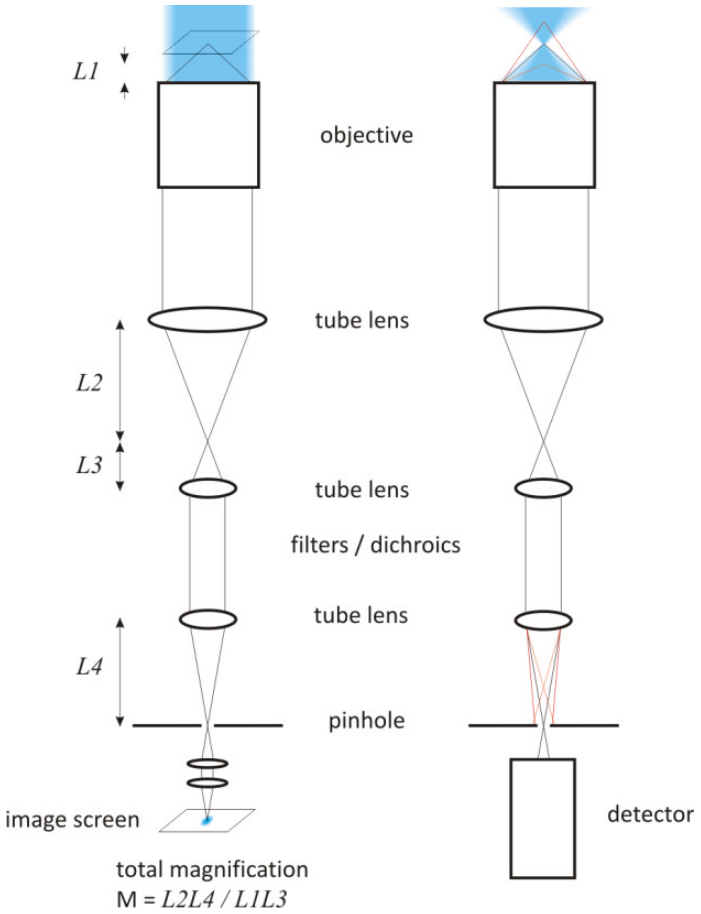

Figure 2.2 **The function of the pinhole under confocal excitation** – (left) All the tube lenses in the optical path determine the total magnification of the sample on the image screen. On the Zeiss ConfoCor2 FCS microscope, a 70-μm pinhole images a 1-Airy unit 488-nm excitation spot created by the C-Apochromat 40x/1.2W objective optimally. This would roughly mean that, for a typical confocal volume diameter of ~500 nm, the total magnification is about 140×. (right) Under confocal excitation, the pinhole acts as an off-plane detection cut-off.

As such, fluorescence in a confocal microscope can be measured in a 'confocal volume' as small as a billionth-of-a-billionth cubic metre, even smaller than a bacterium, as illustrated in some quantitative imaging experiments we performed inside *E. coli* and on the *P. aeruginosa* cell membrane (Figure 2.3). When combined with sample scanning or laser scanning, confocal fluorescence microscopy can generate images of samples with a resolution at the border of what is theoretically possible with visible light, about 200 nm in the radial and 1 μm in the axial direction.

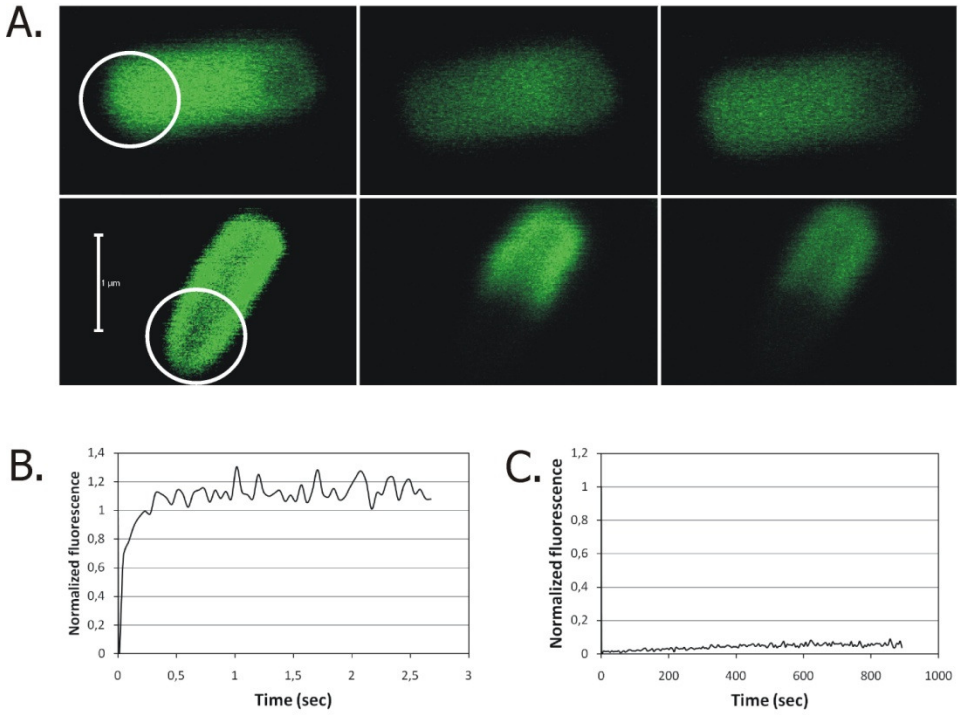

Figure 2.3 **Fluorescence recovery after photobleaching analysis of the peptidoglycan binding domain of *Pseudomonas* phage endolysin KZ144 (PDBKZ-GFP)** (A) An immobilized *E. coli* cell expressing PBDKZ-GFP in its cytoplasm (upper) and an immobilized *P. aeruginosa* cell with bound PBDKZ-GFP molecules (lower) are imaged shortly before photobleaching (left). Subsequently, a pole of an immobilized cell (white circle) is irreversibly photobleached and a new image is collected immediately (middle). After 15 minutes a final image of the cell is collected (right). The scale bar represents 1 µm. (B-C) A time profile of the fluorescence signal within the bleached region is shown. Whereas fluorescence is recovered quickly in the *E. coli* cytoplasm (B), PBDKZ-GFP mobility is low when bound on *P. aeruginosa* peptidoglycan (C). The profiles are average of five (B) or three (C) replicates (Briers et al., 2009).

At nanomolar concentrations of fluorochromes, less than 10 molecules will be present in the confocal volume. With fluorescence correlation and cross-correlation spectroscopy, the fluorescence signal from such low molecule numbers is analysed quantitatively.

2.2. Fluorescence correlation spectroscopy

2.2.1. Introduction

Fluorescence correlation spectroscopy (FCS) is a quantitative fluorescence technique that can be used to determine the concentration and dynamic behaviour of molecules (Magde et al., 1972; Elson and Magde, 1974; Rigler and Elson, 2001). FCS is performed on a confocal microscope where, at low concentrations of fluorochrome, the recorded fluorescence signal spontaneously fluctuates around a constant mean. These fluctuations arise from entry and exit of individual fluorochromes into and out of the confocal volume. Autocorrelation, a mathematical signal processing technique, of the fluctuating fluorescence signal allows to determine both concentration and rate of diffusion. Intuitively it can be understood that the concentration of the fluorochrome determines the amplitude of the fluorescence fluctuations; the higher the concentration, the smaller the amplitude of the fluorescence fluctuations relative to the mean signal. The rate of diffusion on the other hand determines the duration of the fluorescence fluctuations.

2.2.2. The fluorescence autocorrelation function

In signal processing, autocorrelation is used to reveal the similarity of a signal with a time-delayed version of itself, in order to reveal repeating patterns in the signal. For a signal X(t) one can calculate an autocorrelation of time delay τ as follows:

$$r(\tau) = < X(t)X(t+\tau) > \qquad \text{Equation 2.1}$$

Where < > denotes the 'time average', the sum over all products normalized for the elements in the sum. We can rewrite Equation 2.1 to give the normalized fluorescence intensity autocorrelation function I(τ):

$$I(\tau) = \frac{< I(t)I(t+\tau) >}{< I >^2} \qquad \text{Equation 2.2}$$

Usually, fluorescence autocorrelation is defined in terms of the fluorescence fluctuations, δI(t), the deviation from the mean fluorescence intensity <I>:

$$I(t) = < I > + \delta I(t) \qquad \text{Equation 2.3}$$

Now we can define the normalized fluorescence autocorrelation function G(τ) (ACF) (Magde et al., 1972; Elson and Magde, 1974; Rigler and Elson, 2001):

$$G(\tau) = I(\tau) - 1 = \frac{< \delta I(t)\delta I(t+\tau) >}{< I >^2} \qquad \text{Equation 2.4}$$

FCS microscopes generally come with a hard- or software correlator that calculates the ACF in real-time so the experiment can be analysed directly after the measurement. The shape of the ACF is determined by the molecule detection function (MDF) and by the translational and rotational diffusion coefficient and photophysical propeterties of the molecules under observation. The MDF quantifies the efficiency with which a photon is detected from a fluorescing molecule in solution. It depends on the intensity distribution of the focused laser light and on the efficiency of detecting a photon from its point of origin (Schwille et al., 1997; Enderlein et al., 2005). When the MDF is approximated by a Gaussian in axial and lateral directions, the autocorrelation function can be described by an analytical model. In the case of normal Brownian diffusion of fluorescent molecules having no dark states in a homogenous solution under normal excitation conditions, this analytical model takes the following form:

$$G(\tau) = \frac{1}{< N >} G_D(\tau, \tau_{diff})$$

$$G_D(\tau, \tau_{diff}) = \frac{1}{1 + \frac{\tau}{\tau_{diff}}} \sqrt{\frac{1}{1 + \frac{\tau}{S^2 \tau_{diff}}}} \qquad \text{Equation 2.5}$$

With 1/<N> the amplitude of the ACF, τ_{diff} the average diffusion time of the particles and $S=\omega_2/\omega_1$ the structure parameter, which is obtained by calibration. Parameters ω_1 and ω_2 are respectively the radial and axial distances at $1/e^2$ times the maximal emission. Because the latter parameters cannot be obtained independently from fitting an ACF with Equation 2.5, the structure parameter S is used as a fit parameter instead. The amplitude of the autocorrelation function is inversely related with <N>, the average number of molecules in the confocal volume (Figure 2.4). This inverse relationship is logical, the lower the concentration, the bigger the fluctuation relative to the mean signal, so the higher the normalized autocorrelation. The second part of the equation describes the decay of the autocorrelation function. The diffusion time τ_{diff}, the average time a molecule spends in the confocal volume, is defined as the correlation time at which the autocorrelation amplitude has decayed to half its maximal value (Figure 2.4). In Chapter 3 we will investigate which other factors, besides the concentration and diffusion, can affect the shape of the ACF.

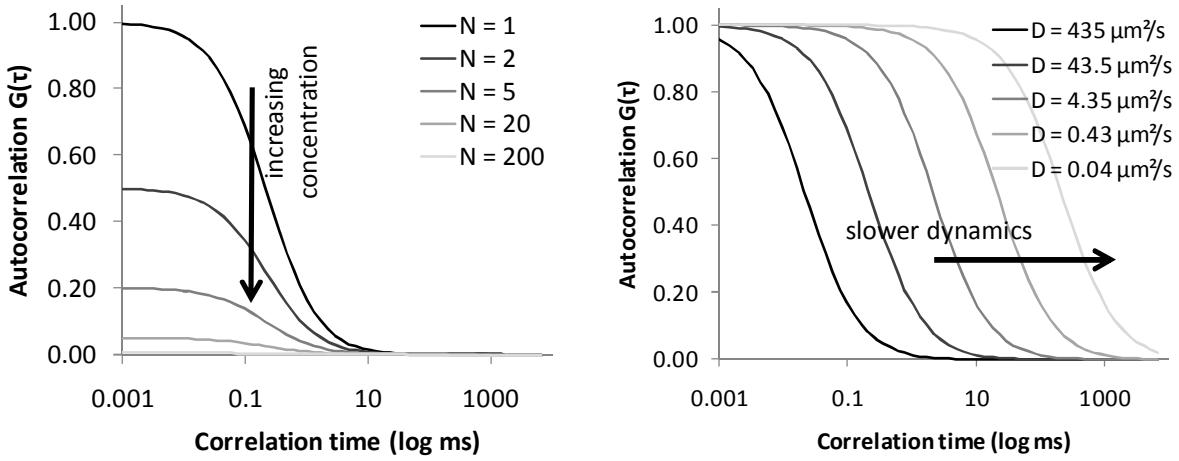

Figure 2.4 **The normalized autocorrelation function** – (left) the higher the concentration, the higher the average number of particles N in the confocal volume, so the lower the autocorrelation function amplitude. (right) the slower the dynamics of the molecule of interest, the smaller the diffusion coefficient of the molecule, the higher the average diffusion time τ_{diff}, so the further the autocorrelation function is shifted to the right.

Generally, the autocorrelation amplitude is a brightness-weighted sum of all the individual species present in the solution (Rigler and Elson, 2001):

$$G(0) = \frac{\sum_i B_i{}^2 N_i}{(\sum_i B_i N_i)^2} \qquad \text{Equation 2.6}$$

Note that in the case of a homogenous solution containing only particles with brightness B, the amplitude indeed simplifies to 1/N. In autocorrelation analysis, the squared dependence on the brightness has important consequences for analysis of mixed species with unequal brightness, as we will discuss further on.

2.2.3. The apparent brightness

The normalized ACF in principle does no longer carry any information about the absolute brightness of molecules, only their brightness relative to each other. However, the ratio of the total average signal <I> with the average number of particles <N>, the counts-per-molecule (CPM), can be used as an 'apparent brightness':

$$CPM\left(\frac{counts}{molec.sec}\right) = \frac{< I >}{< N >} \qquad \text{Equation 2.7}$$

The CPM depends on the 'signal-to-background' ratio of the experiment. More specifically, the 'signal' is determined by the photophysical properties of the fluorophore (extinction coefficient, quantum yield, photophysical states) and by the experimental conditions (laser power, microscope optics, detector quantum yield). The 'background' on the other hand

represents all phenomena that contribute photons to the signal that are not originating from the diffusing molecules present in the 3-dimensional Gaussian detection volume (cellular autofluorescence, out-of-focus emission detection, scatter, APD afterpulsing,...). Importantly, uncorrelated noise will not affect the amplitude of fluorescence fluctuations δI, but it will increase the total average signal <I>. Hence, following Equation 2.4, noise will decrease the normalized ACF amplitude, increase the apparent particle number <N> and consequently, decrease the CPM. For example, for <N> = 4, $<I_{signal}>$=4 kHz and $<I_{noise}>$=4 kHz, the CPM will be 1 kHz/molec in absence of noise, but it will be 0.5 kHz in the presence of noise and moreover, $<N_{apparent}>$ will be 16 instead of 4. The photon counting histogram (PCH) technique (Chen et al., 1999), fluorescence intensity distribution analysis (FIDA) (Kask et al., 1999) or time-integrated fluorescence cumulant analysis (TIFCA) (Wu and Muller, 2005) are alternative techniques to determine the brightness (and stoichiometry) of molecules, although they in principle suffer from the same artefacts and challenges that compromise FCS analysis.

2.2.4. Measuring molecular properties

The motion of a single molecule by diffusion in solution is described by a 'random walk' (Tinoco et al., 2001). At any time t, the distance (d) of the molecule from its original location is the standard deviation σ (Figure 2.5). The variance, σ^2, also called the mean squared displacement (MSD) is given (in two dimensions) by:

$$\sigma^2 = 4Dt \qquad \text{Equation 2.8}$$

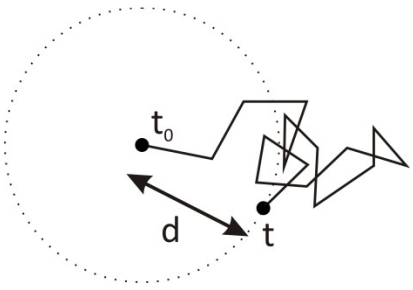

Figure 2.5 **Random walk** – Brownian diffusion of a molecule (black spot) can be approximated by a random walk model, where the molecule takes steps in a random direction with a certain step length. At time t, the distance d from its point of origin is a measure for its dynamic properties

with D the diffusion coefficient, which for a globular molecule is defined by the Stokes-Einstein law:

$$D = \frac{kT}{6\pi\eta r_H}$$ Equation 2.9

with k = 1.38×10^{-23} J/K the Boltzmann constant, T the temperature, η (J.s/m³) the viscosity (pure water 1 cP – 0.001 J.s/m³) and r_H the hydrodynamic radius of the molecule. The diffusion coefficient thus describes the displacement of a molecule in solution over time, given its molecular properties (shape, r_H) and that of the environment (η,T).

In terms of the FCS analysis, measuring τ_{diff} in a focal volume of 'standard deviation' or radius ω_1, allows to calculate an absolute, instrument independent property of the molecule, D:

$$D\ (\mu m^2 \cdot s^{-1}) = \frac{\omega_1^2}{4\tau_{diff}}$$ Equation 2.10

Typically, ω_1 is obtained by measuring the diffusion time of a reference fluorochrome with a known D (Table 2.1).

Table 2.1 **Diffusion coefficient standards** – List of commonly used fluorochromes for FCS calibration with their most accurately determined D. ME = mercaptoethanol derivative, CA = free carboxylic acid, SE = succinimidyl ester derivative.

Fluorochrome	D (µm²/s)	T (°C)	References
rhodamine 6G	426	22.5	(Petrasek and Schwille, 2008)
oregon green	411	25	(Muller et al., 2008)
ATTO 655 ME	407	25	(Muller et al., 2008)
ATTO 655 CA	426	25	(Muller et al., 2008)
alexa fluor 488 SE	435	22.5	(Petrasek and Schwille, 2008)
alexa fluor 546 SE	341	22.5	(Petrasek and Schwille, 2008)
fluorescein	436	22.5	(Paul et al., 1998; Petrasek and Schwille, 2008)
eGFP	95	22.5	(Schenk et al., 2004b; Petrasek and Schwille, 2008)

As said above, D allows to calculate r_H, which for a globular molecule allows to give an estimate about the relative molecular mass, M_r:

$$M_r\ (g \cdot mol^{-1}) = \frac{4}{3}\pi\rho N_A r_H^3$$ Equation 2.11

With ρ the density of the molecule (1.2 g/cm³ for proteins) and N_A Avogadro's number. One can thus see that D and M_r are related as follows:

$$D \sim M_r^{-1/3} \qquad \text{Equation 2.12}$$

Knowing D for a reference protein in a certain medium, allows to estimate the D for another protein in the same medium.

The dimensions of the confocal volume can be determined by measuring τ_{diff} and S for a probe with known D and calculating ω_1 and ω_2. From these parameters the volume can be calculated with:

$$V(m^3) = \pi^{3/2}\omega_1^2\omega_2 \qquad \text{Equation 2.13}$$

Alternatively, the volume can be approximated with a cylinder (applications handbook of the Zeiss ConfoCor2 microscope):

$$V(m^3) = 2\pi\omega_1^2\omega_2 \qquad \text{Equation 2.14}$$

With the volume and the average number of particles known, one can also calculate the absolute concentration of the molecule:

$$C(M) = \frac{< N >}{VN_A} \qquad \text{Equation 2.15}$$

If one assumes a confocal volume of 1 fL (1 μm³), at a concentration of 1 nM about 0.6 particles are on average present in the confocal volume.

2.2.5. Experimental models for FCS

2.2.5.1. Fast fluorochrome intensity fluctuations

When a fluorochrome is in the excited state, there can be a certain probability with which it can enter a short-lived dark state, such as the triplet state, a light-enhanced cis-trans isomerized state, a light induced protonated state, etc. At the excitation intensities used for FCS, fluorochromes frequently enter non-fluorescent states. Triplet conversion for example, renders a fluorochrome dark for a couple of microseconds, and typically appears in the experimental ACF at this time scale.

To account for this in the fitting of the ACF we can expand Equation 2.5 (Widengren et al., 1995; Schwille et al., 2000):

$$G(\tau) = \frac{1}{N}\left[\frac{F_{dark}}{1-F_{dark}}\exp\left(-\frac{\tau}{\tau_{dark}}\right) + G_D(\tau, \tau_{diff,n})\right] \quad \text{Equation 2.16}$$

Mostly, $\tau_{dark} << \tau_{diff}$, and Equation 2.16 is often rewritten as:

$$G(\tau) = \frac{1}{N}\left[\frac{F_{dark}}{1-F_{dark}}\exp\left(-\frac{\tau}{\tau_{dark}}\right) + 1\right] G_D(\tau, \tau_{diff,n}) \quad \text{Equation 2.17}$$

In the presence of a fast dark state, the ACF amplitude is equal to $N(1-F_{dark})$ and thus represents the average number of molecules in the bright state. This model may easily be expanded to contain multiple dark states, as we will see in Chapter 7 when the photophysical properties of red fluorescent proteins are studied.

2.2.5.2. Two diffusing species

In the case a mixture of two different species (different size or brightness) the ACF looks as follows:

$$G(\tau) = \frac{{B_1}^2 N_1}{(B_1N_1+B_2N_2)^2} G_D(\tau, \tau_{diff,1}) + \frac{{B_2}^2 N_2}{(B_1N_1+B_2N_2)^2} G_D(\tau, \tau_{diff,2}) \quad \text{Equation 2.18}$$

with $B_{1,2}$ the brightness of species 1,2 and $N_{1,2}$ the number of particles of species 1,2. In the case that $B_1 = B_2$ this model simplifies to:

$$G(\tau) = \frac{N_1}{(N_1+N_2)^2} G_D(\tau, \tau_{diff,1}) + \frac{N_2}{(N_1+N_2)^2} G_D(\tau, \tau_{diff,2}) \quad \text{Equation 2.19}$$

which in a simpler form looks like:

$$\boldsymbol{G(\tau) = \frac{1}{N_1+N_2}\left[F_1 G_D(\tau, \tau_{diff,1}) + (1-F_1) G_D(\tau, \tau_{diff,2})\right]} \quad \text{Equation 2.20}$$

where total number of particles $N = N_1 + N_2$ defines the amplitude of the ACF, and $F_{1,2}$ are the respective fractions of species 1 and 2. This is the standard model for 2-component diffusion. We can try to rewrite Equation 2.18 similarly:

$$G(\tau) = \frac{1}{N_1+N_2}\left[\frac{B_1^2 F_1}{(B_1F_1+B_2(1-F_1))^2} G_D(\tau, \tau_{diff,1}) + \frac{B_2^2(1-F_1)}{(B_1F_1+B_2(1-F_1))^2} G_D(\tau, \tau_{diff,2})\right] \quad \text{Equation 2.21}$$

Consequently, without preliminary knowledge of B_1 and B_2, fitting with Equation 2.20 works as well for species with unequal brightness, but the fraction components will be scaled by the square of their brightness and by their fraction. Not only that, as we can see from the amplitude of the ACF, the obtained particle number is not the real particle number:

$$G(0) = \frac{B_1^2 N_1 + B_2^2 N_2}{(B_1N_1+B_2N_2)^2} = \frac{1}{(N_1+N_2)\dfrac{(B_1F_1+B_2(1-F_1))^2}{B_1^2F_1 + B_2^2(1-F_1)}} = \frac{1}{N'} \quad \text{Equation 2.22}$$

For example, if $B_2 = 10 \times B_1$ and $N_1 = N_2$, then $F_{2,app} = 100 \times F_{1,app}$ and the apparent particle number is 40% smaller than the real value. Knowledge of the brightness of the particles is thus absolutely necessary for correct concentration determination if particles tend to form oligomers. As we will see further on, Equation 2.22 can also be used to correct for or quantify FRET.

2.2.5.3. Anomalous diffusion

In the case of Brownian diffusion the mean squared displacement of a molecule increases linearly in time (Equation 2.8). Under circumstances of anomalous or obstructed diffusion, at short time scales diffusion still seems normal but as time increases the MSD seems to lag behind (Crank, 1975). Empirically, under these circumstances the MSD follows a power law of t:

$$MSD = 4Dt^{\alpha} \qquad \text{Equation 2.23}$$

With $\alpha<1$ the anomaly parameter describing the strength of obstruction (Figure 2.6 left). We can modify Equation 2.5 to take anomalous diffusion into account (Schwille et al., 1999; Wachsmuth et al., 2000):

$$G(\tau) = \frac{1}{< N >} \frac{1}{1 + \left(\frac{\tau}{\tau_{diff}}\right)^{\alpha}} \sqrt{\frac{1}{1 + \frac{1}{S^2}\left(\frac{\tau}{\tau_{diff}}\right)^{\alpha}}} \qquad \text{Equation 2.24}$$

Since the diffusion coefficient has units of $\mu m^2/s^{\alpha}$, it can no longer be interpreted in terms of molecular size, even though the correlation time at half amplitude is unaffected by an anomalous trend of the ACF (Figure 2.6 right). Apparent anomalous diffusion can be a consequence of obstructed diffusion (Saxton, 1994; Weiss et al., 2004), or binding to immobile structure (Saxton, 1996), as has been simulated for chromatin binding (Wedemeier et al., 2008) or in some cases heterogeneous particle size distributions. Binding of proteins to chromatin will be discussed in detail in Chapter 4.

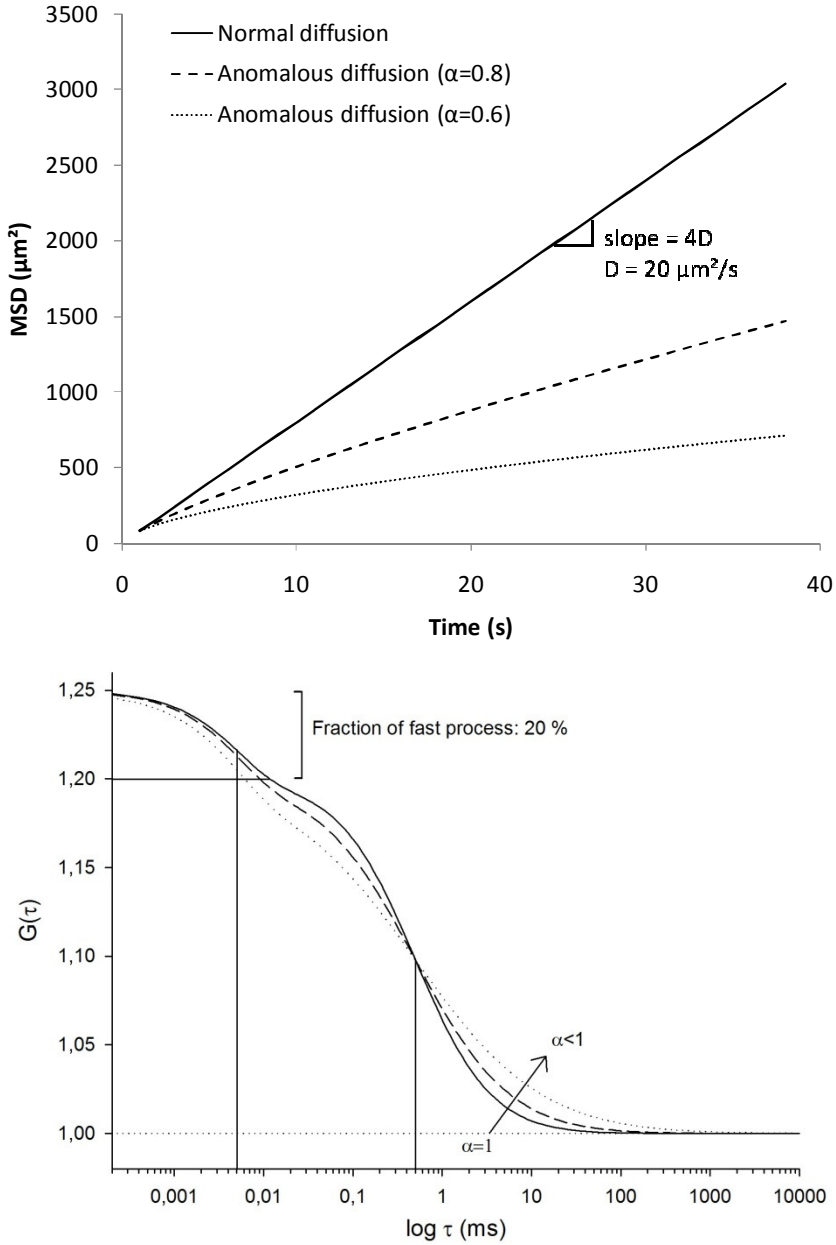

Figure 2.6 **Anomalous diffusion** – (top) Mean square displacement versus time for normal, anomalous and superdiffusion (Crank, 1975). (bottom) Simulated ACFs in a confocal volume with S=5, N=5 and a diffusion time of 500 μs. The relaxation time of the fast process is 5 μs. The anomaly parameter α is 1,0 (——); 0,8 (- - -) en 0,6 (·····).

2.2.5.4. Rotational diffusion

Diffusing molecules move both by translational and rotational diffusion. Next to translational diffusion, rotational diffusion, in principle, is also observed in the ACF of a fluorochrome when using linearly polarized excitation light. In eGFP for example, the rigid fluorophore is part of the protein, so rotational diffusion of eGFP can be observed in the tens of nanoseconds time scale. Generally, when the excitation and emission polarization directions

are parallel, a high autocorrelation is observed at a time scale faster than the rotational correlation time. When excitation and emission polarization directions are perpendicular, an anti-correlation is seen at this time scale. Since the time scales of rotational and translational diffusion are usually well separated, the autocorrelation function can be expanded with an exponential to take the rotational diffusion into account (Ehrenberg and Rigler, 1974; Kask et al., 1987; Widengren et al., 1999a):

$$G(\tau) = 1 + \frac{1}{< N >}\left[1 + A\exp(-\frac{\tau}{\theta})\right] G_D(\tau) \qquad \text{Equation 2.25}$$

With A, the amplitude of the exponential, that depends on the geometry of the experimental setup and θ the experimental rotational correlation time. Rotational diffusion is, due to its third-power dependence on the hydrodynamic radius of the molecule, much more sensitive to study the size of particles:

$$\theta = (6D_r)^{-1} = \frac{4\pi\eta r_H^3}{3RT} \qquad \text{Equation 2.26}$$

With D_r the rotational diffusion coefficient and R = 8.314 472 JK^{-1} mol^{-1} the gas constant.

2.2.6. Quenching

FCS is often used in ligand binding studies. It can happen that the brightness of the fluorochrome is altered upon binding. By introducing a parameter R, the relative brightness of the fluorochrome in the bound versus the free state (Foldes-Papp, 2005), this quenching can be taken into account:

$$G(\tau) = \frac{N_{free} + R^2 N_{bound}}{\left(N_{free} + RN_{bound}\right)^2} \qquad \text{Equation 2.27}$$

Quenching will result in a lower apparent brightness. For fluorescent proteins, quenching is often not a problem since the fluorophore lies inside of the rigid beta-barrel structure and is rather insensitive to the environment.

2.2.7. Advanced FCS

In its original form, FCS is performed on a setup with a continuous-wave excitation laser and a single channel detection unit. Although in many cases this is adequate, continuing efforts are made to improve the performance of FCS. These efforts, illustrated in Figure 2.7 and discussed in the following sections, can be divided into 4 categories: increasing the range of concentrations, increasing the timescale of dynamics, increasing the statistical accuracy and precision and increasing the spatial information per measurement.

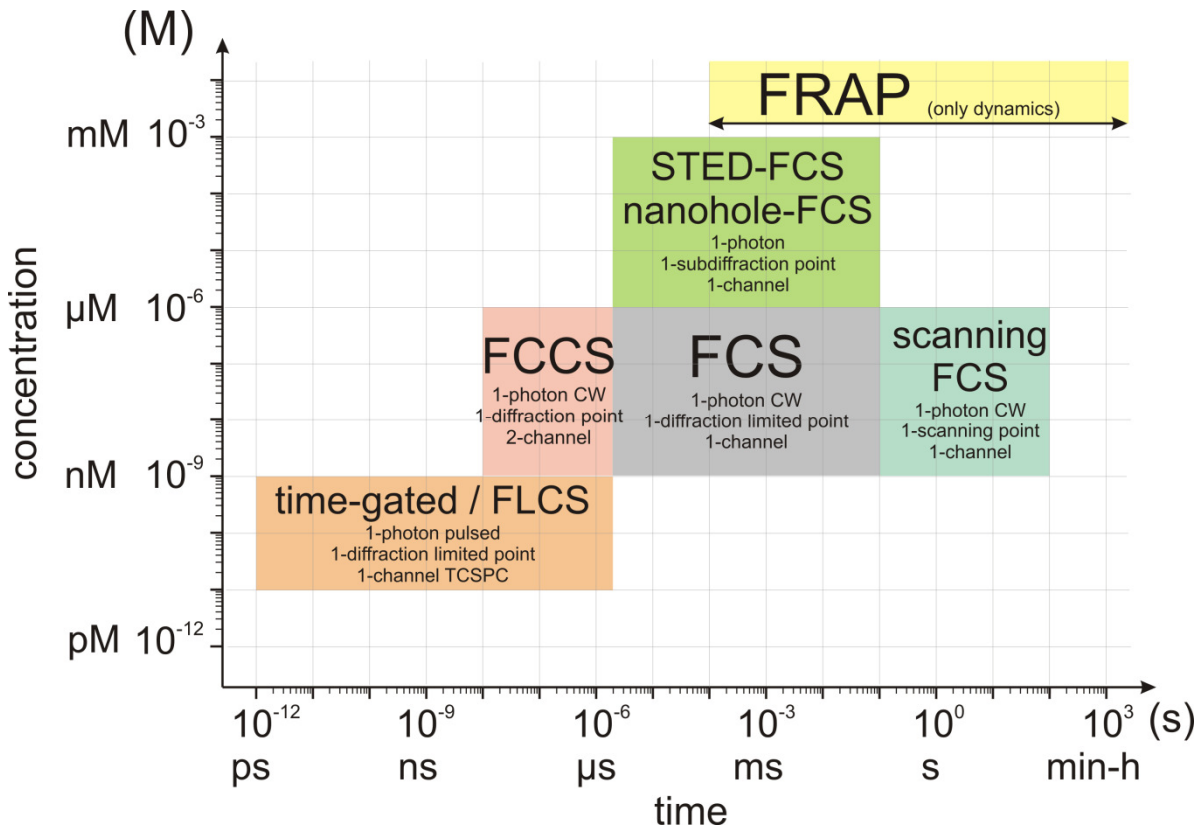

Figure 2.7 **Time scale and dynamic range of different FCS based methods** – For details on the techniques see sections 2.2.7.1-2.2.7.5.

2.2.7.1. Low concentrations and fast dynamics

The signal-to-noise ratio of an FCS experiment, as discussed in 2.2.3, sets the lower limit for accurate concentration measurements to about 1 nM. By using picosecond pulsed lasers instead of CW-lasers in combination with time-correlated single photon counting (TCSPC) instrumentation, this barrier can be overcome. With time-gated FCS, only photons arriving in a certain 'time gate' with respect to the excitation pulse are used for calculating the ACF (Lamb et al., 2000). Like this, the signal originating from background or scatter can be greatly suppressed. With time-resolved FCS, also known as fluorescence lifetime correlation spectroscopy (FLCS), the fluorescence decay curve of the fluorochrome is used to calculate a 'statistical filter', that in term is used to eliminate all non-wanted photons from the signal prior to the calculation of the ACF, effectively eliminating the noise (Bohmer et al., 2002; Kapusta et al., 2007). Both with time-gated FCS as with FLCS, the accuracy of the ACF at extremely low concentrations is greatly improved and FCS measurements can be performed reliably down to the pM range. With time-gated FCS and FLCS the ACF can be measured from the picosecond time scale all the way to the standard FCS timescale, eliminating the

need for a dual detector cross-correlation setup to avoid detector artefacts (Widengren et al., 1995; Enderlein and Gregor, 2005). In this way, ultrafast dynamics, unaccessible by standard FCS can now be studied, such as rotational diffusion, as discussed in section 2.2.5.4. Also, the antibunching properties of the fluorochrome can be measured, and stoichiometries can be determined (Sykora et al., 2007).

2.2.7.2. High concentrations

At high concentrations the low fluctuation amplitude decreases the precision of an FCS measurement. This sets an upper limit of about 1 μM to the concentrations that can be measured accurately with FCS. A straightforward solution to this problem would be to decrease the confocal volume in order to increase the fluctuation amplitudes. Some methods have been put forward to circumvent the diffraction-limitation of FCS measurements. The first method is based on total internal reflection and allows to measure the concentration of molecules close to or bound to the glass surface (Thompson et al., 1981). The second method is based on stimulated emission depletion of the fluorochromes everywhere but in the centroid of the focal spot, leading to an apparent smaller focal spot (STED-FCS) (Hell and Wichmann, 1994; Eggeling et al., 2009). The last method is based on the physical limitation of the focal spot in solution by measuring on a cover glass with a metal film containing nanoholes (Levene et al., 2003).

2.2.7.3. Slow dynamics

Fluorescence fluctuations arise from the diffusion of molecules. If a protein or protein complex is very slow or immobile, fluorescence will not fluctuate fast enough or not at all for FCS to accurately measure protein dynamics and interaction. In addition, because FCS is a point-measurement, the same fluorochromes will be illuminated the whole time. This will lead to photobleaching, the irreversible loss of fluorochrome fluorescence. If photobleaching is significant, the apparent diffusion time will decrease and the apparent particle number will continuously decrease. Photobleaching of slow molecules can be avoided by keeping the excitation stress low, while still allowing to acquire as many photons as needed for a statistically good measurement. Originally, this was done by physically moving the sample with respect to the laser (Petersen, 1986), but nowadays laser scanning has largely replaced sample scanning. In the line scanning FCS and circle scanning FCS method, the laser is moved in discrete patterns across the sample (Meyer and Schindler, 1988; Berland et al., 1996; Ruan et al., 2004; Petrasek and Schwille, 2008; Ries et al., 2009), allowing to study slow

dynamics more accurately. In the image correlation spectroscopy (ICS) method, the size and concentration of immobile clusters can be determined (Petersen et al., 1993). The raster image correlation spectroscopy (RICS) method allows to calculate both fast and slow dynamics by a spatio-temporal correlation of individual pixels in a LSM image (Digman et al., 2005). What all the above cited methods have in common is that the exposure time of the molecules can be tuned to their mobility, thus avoiding photobleaching. Finally, a mathematical 'trick' also exists to correct for photobleaching by correcting the count rate trace prior to autocorrelation (Ries et al., 2009):

$$F_{corrected}(t) = \frac{F_{recorded}(t)}{\sqrt{f(t)/f(0)}} + f(0)\left(1 - \sqrt{f(t)/f(0)}\right) \qquad \text{Equation 2.28}$$

Where f(t) is a function used to fit the experimental count rate trace

$$f(t) = y0 + \sum_i A_i \exp\left(-t/t_0\right) \qquad \text{Equation 2.29}$$

Correcting the ACF in this way does restore the amplitude and shape of the ACF when photobleaching occurs, but it does not correct or quantify the apparent faster diffusion time.

2.2.7.4. Absolute diffusion coefficients

Determining absolute diffusion coefficients with FCS can be challenging, since the exact shape of the confocal volume is hard to take into account when analysing the ACF. Several FCS methods that use an extrinsic ruler to determine absolute diffusion coefficients have recently been proposed. Enderlein reported two-focus or dual-focus FCS (Dertinger et al., 2007b). In this method a Nomarski prism is placed in front of the back aperture of the objective. When two coinciding alternatively pulsing excitation (pulsed interleaved excitation, PIE) beams, each linearly polarized but orthogonal to each other, pass through this prism, a small ~200-nm shift occurs between them. These two beams are focussed by the objective and two focal spots are created next to each other. Emission from these spots is finally detected in cross-correlation mode. The diffusion time from the cross-correlation function of one spot with the other is a function of only the properties of the Nomarski prism. All optical artefacts (cover glass thickness, astigmatism, refractive index) are cancelled out by this way of measuring, providing a way to measure accurately the absolute diffusion coefficient (Table 2.1). In addition, 2-focus FCS seems to be more precise than standard FCS, allowing to measure differences as small as 3% in D (Dertinger et al., 2007a). Furthermore, since 2-

focus FCS is performed on a pulsed laser setup, there is an extra advantage of removing the noise. Another method for determining absolute diffusion coefficients is dual-line scanning FCS, because the distance between two arbitrary scan lines is known with very high accuracy (Ries and Schwille, 2006). Finally, circle scanning FCS can also be used for absolute concentration measurements if the circle diameter is calibrated well (Petrasek and Schwille, 2008).

2.2.7.5. Spatial information

Standard FCS measurements are point measurements. Spatial information can logically be obtained by measuring consecutively at different spots. Recently however, a microscopy setup was proposed that uses two independent foci, allowing the simultaneous measurement of FCS (or FCCS) in two arbitrary points (Ferrand et al., 2009). Determining (slow) dynamics in whole images at once is possible with temporal image correlation spectroscopy (ICS) and variants thereof (spatiotemporal ICS, k-space ICS)(Hebert et al., 2005; Kolin et al., 2006). This holds great promise for the future of correlation analysis in biological samples. Very recently a method, based on second-order correlation functions of photoblinking fluorophores, has been put forward, super-resolution optical fluctuation imaging (SOFI), that can increase the resolution of wide-field imaging in three dimensions (Dertinger et al., 2009).

2.3. Fluorescence cross-correlation spectroscopy

2.3.1. Introduction

Dual-color or two-color fluorescence cross-correlation spectroscopy (FCCS) is an extension of FCS that allows studying the interaction between molecules (Schwille et al., 1997). The principle of this technique, fluorescence fluctuation measurements, is similar to that of FCS, but in FCCS two spectrally separated excitation and detection paths allow independent studying of two fluorochromes. More specifically, autocorrelation of the two signals provides dynamic and concentration information about each of the molecules and cross-correlation of the two signals provides information on the interaction between the molecules. Intuitively this can be understood as follows: if a complex of both molecules crosses the confocal volume then two independent fluorescence peaks will be registered at the same time. However, if the molecules are not interacting the two fluorescence peaks will likely not coincide in time.

2.3.2. The fluorescence cross-correlation function

In signal processing, cross-correlation is used to reveal the similarity of two signals as a function of the time delay between them. For two signal X(t) and Y(t), one can calculate the normalized cross-correlation as follows:

$$r(\tau) = < X(t)Y(t+\tau) > \qquad \text{Equation 2.30}$$

Similar as for FCS, we can write for the normalized fluorescence cross-correlation function $G_{CC}(t)$ (Schwille et al., 1997; Rigler et al., 1998):

$$G_{CC}(\tau) = \frac{< \delta I_g(t)\delta I_r(t+\tau) >}{< I_g >< I_r >} \qquad \text{Equation 2.31}$$

Where I_g and I_r represent the average fluorescence intensity in the green and red detection channel, respectively. Generally, as for autocorrelation analysis, the cross-correlation amplitude is a weighted sum of the individual species (Schwille, 2001):

$$G_{CC}(0) = \frac{\sum_i B_{i,g}B_{i,r}\,N_i}{\left(\sum_i B_{i,g}N_i\right)\left(\sum_i B_{i,r}N_i\right)} = \frac{1}{< N_{CC} >} \qquad \text{Equation 2.32}$$

Where N_i is a particle with brightness $B_{i,g}$ in the green channel and $B_{i,r}$ in the red channel. The mathematical model describing the whole CCF is similar as the one given in Equation 2.5, with the amplitude '1/<N>' from Equation 2.32. We can simplify Equation 2.32 for the simplest case of a mixture of monomeric and heterodimeric particles:

$$\begin{aligned} G_{CC}(0) &= \frac{B_g B_r N_{gr}}{B_g(N_g + N_{gr})B_r(N_r + N_{gr})} \\ &= \frac{N_{gr}}{(N_g + N_{gr})(N_r + N_{gr})} = \frac{N_{gr}}{N_{g,total}N_{r,total}} \end{aligned} \qquad \text{Equation 2.33}$$

This is the well-known standard formula for FCCS analysis (Schwille et al., 1997; Rigler et al., 1998; Schwille and Haustein, 2001; Schwille, 2001; Foldes-Papp, 2005). This amplitude is inversely proportional to the total concentration of both green and red particles, similarly to an FCS measurement, and directly proportional to the number of double-labeled particles. When proteins do not interact, N_{gr} is 0, so the cross-correlation will be 0, the higher N_{gr}, the higher the cross-correlation. When the solution contains only heterodimeric particles, or if $N_{gr} >> N_g, N_r$, i.e. almost all particles are double labelled, then the cross-correlation amplitude is inversely related to the N_{gr} (Rigler et al., 1998):

$$G_{CC}(\tau) = \frac{1}{N_{gr}} \qquad \text{Equation 2.34}$$

When the green (or red) protein is in excess, then Equation 2.33 simplifies to:

$$G_{CC}(\tau) = \frac{1}{N_{g\ (or\ r)}} \qquad \text{Equation 2.35}$$

When both proteins are in excess over their complex, Equation 2.33 simplifies to (Rigler et al., 1998):

$$G_{CC}(\tau) = \frac{N_{gr}}{N_g N_r} \qquad \text{Equation 2.36}$$

2.3.3. Measuring molecular properties

Since the two optical paths of the two fluorochromes generally have different molecular detection functions, the confocal volumes V_{green} and V_{red} often are not the same. In fact (discussed further in section 2.3.4), the red volume is always slightly larger than the green volume. We can approximate the cross-correlation volume, where double labeled particles can be detected, with the root mean square of the axial and radial radii of the green and red confocal volume (Schwille et al., 1997):

$$\omega_{1,CC} = \sqrt{\frac{\omega_{1,G}^2 + \omega_{1,R}^2}{2}} \qquad \text{Equation 2.37}$$

$$\omega_{2,CC} = \sqrt{\frac{\omega_{2,G}^2 + \omega_{2,R}^2}{2}} \qquad \text{Equation 2.38}$$

$$V_{CC} = \left(\frac{\pi}{2}\right)^{3/2}\left(\omega_{1,g}^2 + \omega_{1,r}^2\right)\sqrt{\left(\omega_{2,g}^2 + \omega_{2,r}^2\right)} \qquad \text{Equation 2.39}$$

This volume lies in between the green and red volumes. The diffusion coefficient of the double labelled particles, by analogy with Equation 2.10, is defined as (Schwille et al., 1997):

$$D\ (\mu m^2 \cdot s^{-1}) = \frac{\omega_{1,G}^2 + \omega_{1,R}^2}{8\tau_{diff}} \qquad \text{Equation 2.40}$$

The concentration of double labelled particles can be calculated by analogy with Equation 2.15:

$$[GR](M) = \frac{N_{gr}}{V_{CC} N_A} \qquad \text{Equation 2.41}$$

N_{gr} can be obtained from Equation 2.33 simply by measuring the amplitudes of the ACFs and CCF and V_{CC} can be obtained from Equation 2.39.

2.3.4. Non-idealities

Under ideal experimental conditions for FCCS, the two fluorochromes used are each observed independent from each other in an identical confocal volume in solution. In practice however, there are some critical non-idealities that need careful consideration, if quantitative analysis of FCCS is needed. In Chapter 6, we will take these non-idealities into account when we analyse an FCCS experiment quantitatively. In the following sections, the common non-idealities volume overlap, crosstalk, FRET and dark states will be introduced.

2.3.4.1. Volume overlap

Ideally green and red molecules are 'seen' in the same confocal volume. In general terms, this means the two excitation and emission paths image the very same spot in solution. This has two requirements. First, excitation spots for different colors should overlap both in XY as in Z. Second, the two fluorochromes should have similar molecular detection efficiencies in their excitation spot. Although the second requirement depends on the choice of fluorochromes and on the calibration of the microscope, the first requirement is less trivial to fulfill. On a commercial multi-laser setup, lasers are coupled into a single optical fiber which is coupled into the microscope. This allows independent calibration of the laser coupling unit and the microscope and moreover, it makes different laser lines create a concentric multicolour beam in the microscope. However, because of the wavelength dependence of optical refraction, most objectives focus light of a different color in a different spot, a phenomenon called axial chromatic aberration (Figure 2.8).

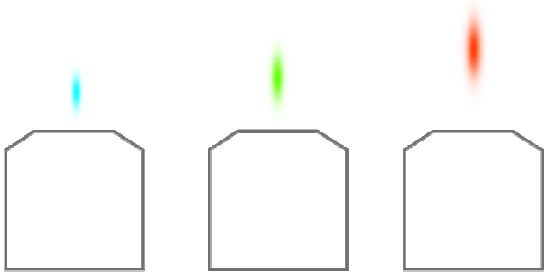

Figure 2.8 **Axial chromatic aberration** The more red the excitation wavelength is, the bigger the excitation spot will be and the farther it is located from the objective front aperture.

Even objectives that are maximally corrected for optical aberrations, 'Apochromats', still exhibit axial chromatic aberration, which is especially noticeable when two wavelengths differ much, for example 488/633-nm excitation. In what follows, the parameters from an FCCS

measurement are first evaluated in the presence of incomplete overlap, after which some methods will be discussed that allow to minimize or avoid incomplete volume overlap.

2.3.4.1.1. Evaluation

If the sizes of the different volumes (green, red and cross-correlation volume) are known, then an incomplete volume overlap can be taken into account when analyzing the curves. We illustrated the effect of and incomplete overlap in Figure 2.9. The particle numbers in Equation 2.33 are each observed in their respective confocal volume (Schwille et al., 1997; Rigler et al., 1998; Bacia et al., 2002; Kohl et al., 2005; Oyama et al., 2006):

$$G_{CC}(0) = \frac{\overset{V_{CC}}{\uparrow}\ N_{gr}}{(N_{gr} + N_g)(N_{gr} + N_r} \qquad \text{Equation 2.42}$$
$$\underset{V_G}{\downarrow} \qquad \underset{V_R}{\downarrow}$$

2.3.4.1.2. Optimizing the pinhole diameters

In principle, by opening the green pinhole and closing the red pinhole a little, the effective size of the confocal volumes can be changed (Schwille et al., 1997). First, one calibrates the XYZ positions of the pinholes at the two respective excitation wavelengths (for example 488 nm for the green detection path and 543 nm for the red detection path). The pinholes then image their conjugate excitation spots perfectly. In a next step, an FCCS measurement is performed with only one laser on a probe that emits in the two detection channels. If the two pinholes image exactly the same spot in solution, the three curves (one CCF and two ACFs) overlap completely. If not, the pinholes can be tuned such that the amplitudes of the three curves overlap. This step can be repeated with the other laser and should provide the same overlap. Importantly, the particle numbers obtained at both excitation wavelengths need not be the same, considering the wavelength-dependence of refraction and considering possible excitation intensity dependent photophysics such as optical saturation (section 3.3.5) and dark state formation (section 7.3.1.1). Finally, when two excitation wavelengths are used that give focal spots with a significant axial shift, such as 488 nm and 633 nm, pinhole optimization cannot completely correct for incomplete overlap. It is then better to optimize the beam diameter.

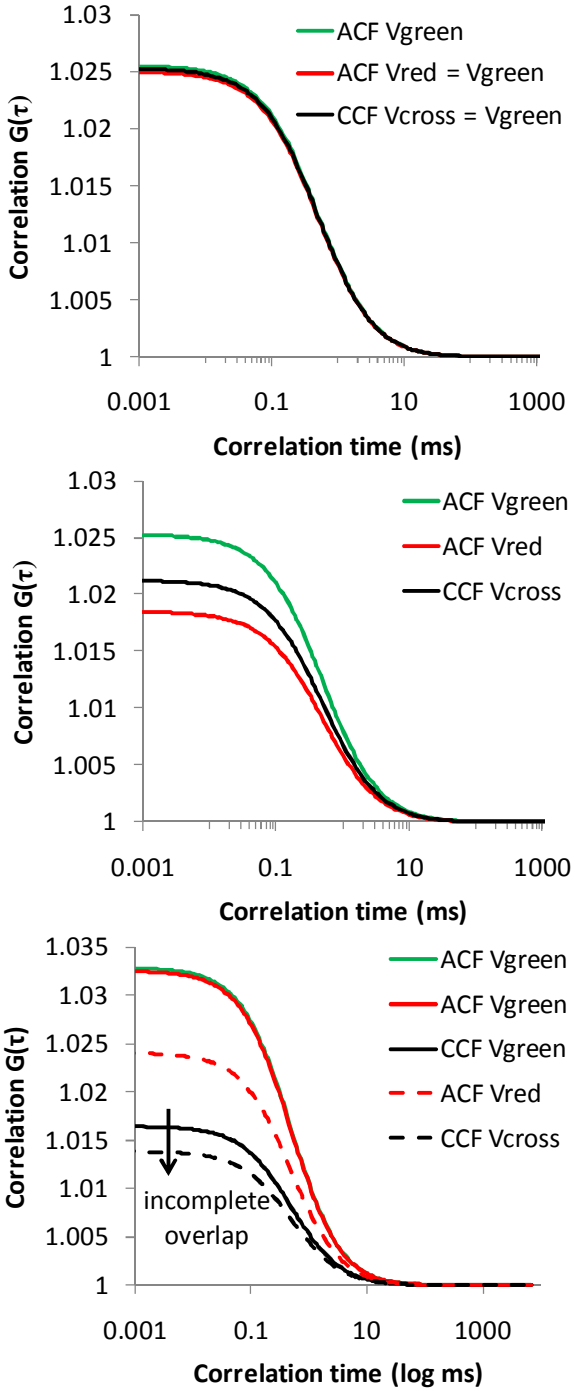

Figure 2.9 **Incomplete overlap** (top and middle) Simulated ACFs and CCF for the case of a solution containing only green-and-red particles and green:red volume ratio of 1:1 (top) or 1:1.37 (middle) as is the case for the Zeiss ConfoCor2 microscope. The particle numbers obtained from each CF scale with the size of the respective volume. (bottom) Simulated ACFs and CCF for the case of a solution 100 nM green-and-red particles and 100 nM green-only and red-only particles. Incomplete overlap decreases the CCF amplitude.

2.3.4.1.3. Optimizing the beam diameter

By decreasing the diameter of the laser beam entering the back aperture of the objective, the effective numerical aperture of the objective at that wavelength is reduced, resulting in a slightly larger focal point. Like this, the overlap of the two focal excitation points can be optimized (Figure 2.10). On a commercial microscope, this is usually not possible since laser lines are combined in a single fiber before entering the microscope.

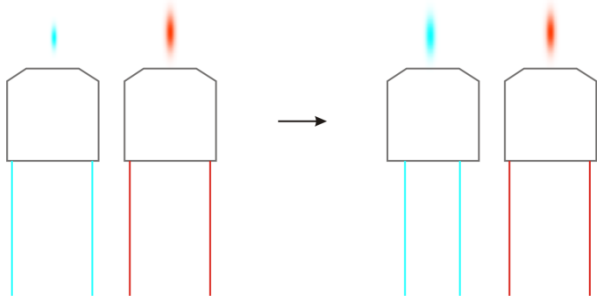

Figure 2.10 **Optimizing beam diameter** – By decreasing the blue beam diameter, a bigger blue excitation spot is created that overlaps better with the red excitation spot.

To calibrate such a setup, a fluorochrome is needed that is excitable by the two lasers and emits in the two detection channels. Absolute (optical saturation error free and dark state error free) particle numbers are obtained by extrapolation to zero excitation intensity. If the ACF amplitude is independent of the laser used to excite the probe, or the channel in which it is detected, then a perfect 3-dimensional volume overlap is achieved.

2.3.4.1.4. Single wavelength or two-photon excitation

When only a single wavelength is used to excite the two fluorochromes (SW-FCCS), all auto- and cross-correlation functions are measured in the same volume, provided that the pinholes are correctly aligned (see section 2.3.4.1.2). For example, it has been shown that by using only the 514-nm Ar-ion laser line both eGFP and mRFP can be excited at the cost of a lower CPM for both fluorochromes (Hwang and Wohland, 2004). Similarly, by employing cyanFP in combination with a fluorescent protein with an ultra-long stokes shift (mKeima), SW-FCCS could successfully be performed at 458 nm excitation (Kogure et al., 2006). Finally, it has been shown that eGFP and Alexa633 can be both efficiently excited at 920-nm by two-photon excitation (Kim et al., 2005), avoiding an incomplete volume overlap and more importantly, significantly reducing cross-talk (which will be discussed in the next section).

2.3.4.2. Cross-talk

Ideally, the fluorescence of the green and red fluorochrome can be spectrally separated in the detection part of the microscope but in reality there is always some unwanted emission of the green fluorochrome into the red detection channel, called cross-talk. One can minimize cross-talk by choosing a fluorochrome pair with large spectral separation, but in cells one is very often obliged to using fluorescent proteins, which have a significant spectral overlap.

2.3.4.2.1. Evaluation

Because cross-talk cross-correlates with the signal in the green channel, it can pose a problem for quantitative FCCS measurements. Suppose one is working at equimolar concentrations of green and red protein and the emission of the green protein in the red channel is 10 times less than the red protein in the red channel. We can account for cross-talk of the green fluorochrome in the cross-correlation amplitude (Equation 2.32):

$$G_{CC}(\tau) = \frac{B_g B_c N_g + 0 B_r N_r + B_g (B_r + B_c) N_{gr}}{(B_g N_g + 0 N_r + B_g N_{gr})(B_c N_g + B_r N_r + (B_r + B_c) N_{gr})} \qquad \text{Equation 2.43}$$

With B_g the brightness of the green fluorochrome in the green channel, B_r the brightness of the red fluorochrome in the red channel and B_c the brightness of the green fluorochrome in the red channel. Dividing numerator and denominator with $B_g B_r$ simplifies the expression (Rigler et al., 1998; Foldes-Papp, 2005):

$$\begin{aligned} G_{CC}(\tau) &= \frac{N_{gr} + (N_g + N_{gr})Q}{(N_g + N_{gr})\left[N_r + N_{gr} + (N_g + N_{gr})Q\right]} \\ &= \frac{N_{gr} + N_{g,total} Q}{N_{g,total}(N_{r,total} + N_{g,total} Q)} \end{aligned} \qquad \text{Equation 2.44}$$

with

$$Q = \frac{B_c}{B_r} = \frac{the\ red\ emission\ of\ the\ green\ fluorochrome}{the\ red\ emission\ of\ red\ fluorochrome} \qquad \text{Equation 2.45}$$

To calculate N_{gr}, one has to know $\langle N_{cc} \rangle$, Q, $N_{g,total}$, $N_{r,total}$. Q can be measured in an *in vitro* experiment, $\langle N_{cc} \rangle$ is measured in the CCF and $N_{g,total}$ is measured in the green ACF. $N_{r,total}$ can only be known from a single color excitation red AC measurement, since in two color excitation Equation 2.21 would have to be used to fit the data, where N_{total} and F_1 would be both unknown. ($F_1 = N_{red}/N_{total}$). Importantly, when $N_{gr} >> N_g$ and N_r, then Equation 2.44 simplifies to $1/N_{gr}$, i.e. cross-talk can be ignored. Likewise, for a mixture of green and green-and-red molecules, $N_r = 0$, and the CCF amplitude simplifies to $1/(N_g+N_{gr})$. For fluorescent

proteins, Q can be measured *in vitro* in a viscous buffer, to separate photophysics and diffusion well. Equimolar concentrations of the proteins are mixed, as judged by the amplitude of the ACFs measured in single color excitation mode. The signal in the red channel upon 488 excitation relative to dual color excitation is equal to Q. When using mRFP1/mCherry and eGFP as fluorescent tags Q ≈ 0.2. We will discuss this further in Chapter 6 when we analyse FCCS of mRFP-eGFP and mCherry-eGFP in detail. Finally, as we have discussed in section 2.2.5.2 the dim green particles observed in the red channel will not contribute significantly to the decay of the red FCS curve ($F_{dimgreen}=0.01F_{red}$), but the red particle number will be underestimated by 40%, which has to be corrected for.

2.3.4.2.2. FCCS without cross-talk

A disadvantage of fluorescent proteins is their limited spectral range, which gives rise to both cross-talk (2.3.4.2.1) and FRET (as we will discuss in the next section), when performing FCCS as originally proposed (Schwille et al., 1997). Although in principle possible to correct for, they complicate FCCS analysis considerably. Over the last years, variations on the original FCCS method have emerged that rely on the temporal separation of fluorochrome excitation, avoiding cross-talk completely. Thews et al reported a method for the temporal separation of excitation of eCFP and eYFP with detection on one detector, by using acousto-optical modulated CW-excitation at 425-nm and 515-nm (Thews et al., 2005). This method, also called alternating laser excitation (ALEX) in FRET studies, completely avoids emission cross-talk. Muller et al concurrently reported a method, pulsed interleaved excitation (PIE) FCCS, with both temporal excitation separation as spectral emission separation by using a dual detection setup and pulsed diode lasers (Muller et al., 2005). This setup has as additional advantage that APD afterpulsing and excitation cross-talk is also completely avoided. Finally, Takahashi et al. introduced a simple method to avoid cross-talk on a microscope with CW-lasers by using an acousto-optical tunable filter in the excitation path (switching method) (Takahashi et al., 2008). The availability of a 531-nm pulsed diode has recently opened the way for exciting both GFPs and RFPs with picosecond pulsed diodes, allowing background and scatter light to be removed in addition to all previously mentioned advantages, which increases the S/N considerably.

2.3.4.3. FRET

Förster resonance energy transfer (FRET) is the radiationless transfer of the excited state energy of a donor fluorochrome to an acceptor fluorochrome, provided that the donor emission spectrum overlaps with the acceptor excitation spectrum and provided that the distance between both fluorochromes is not too large (generally smaller than 10 nm). The efficiency of this process is expressed with the FRET efficiency E, which varies between 0 (no energy transfer) and 1 (complete transfer). The intermolecular distance has a strong effect on the FRET efficiency, due to the dipole-dipole coupling mechanism in FRET. The intermolecular (interfluorochrome) distance r can be estimated from E with:

$$r = R_0 \sqrt[6]{\frac{(1-E)}{E}} \qquad \text{Equation 2.46}$$

With R_0 the Förster radius of the fluorochrome couple, the distance between the fluorochromes when the FRET efficiency has decreased by 50%. How E is measured will be discussed in Chapter 5. We compared the different fluorescent protein couples generally used for FCCS (Table 2.2). Since these fluorescent protein couples perform quite well in FRET, as indicated by the small R_0 values, we evaluated the effect of FRET on the obtained parameters from FCCS.

Table 2.2 **Comparison of different fluorescent protein couples used for FCCS** – R_0 = Förster distance, $\Delta\lambda_{em,max}$ = wavelength difference in the fluorescence emission maxima of the two fluorochromes, a larger difference is beneficial for FCCS, since it lowers cross-talk, photostability and brightness are represented as + (good) or – (not good) both for the green as for the red partner. Finally, the performance of the couples in FRET and FCCS are given.

couple	R_0	$\Delta\lambda_{em,max}$	Photostability	Brightness	Reference	FRET	FCCS
eGFP/mRFP1	4.7	100	+/-	+/+	(Campbell et al., 2002)	++	++
eGFP/mCherry	5.1	103	+/+	+/+	(Shaner et al., 2004)	++	++
eGFP/TagRFP-T	~5	77	+/++	+/++	(Shaner et al., 2008)	++?	+++?
eCFP/mKeima	-	144	+/+	+/-	(Kogure et al., 2006)	-	+

In principle, a reduction of the green fluorochrome brightness and an increase of the red fluorochrome brightness is observed, when the photophysical properties of the fluorochromes and the interfluorochrome distance and orientation allow for FRET. The amplitude of the ACFs can be calculated with Equation 2.22:

Green ACF

$$\mathrm{N_{g,app}} = \frac{\left(N_g + (1-E)N_{gr}\right)^2}{N_g + (1-E)^2 N_{gr}} = N_{g,real}\frac{\left(F_g + (1-E)F_{gr}\right)^2}{F_g + (1-E)^2 F_{gr}}$$ Equation 2.47

Red ACF

$$\mathrm{N_{r,app}} = \frac{\left(N_r + (1-E)N_{gr}\right)^2}{N_r + (1-E)^2 N_{gr}} = N_{r,real}\frac{\left(F_r + (1+E)F_{gr}\right)^2}{F_r + (1+E)^2 F_{gr}}$$ Equation 2.48

Knowing the fractions of green-only, red-only and green-and-red molecules, these equations allow to directly calculate the real particle numbers from the ACF amplitude. The amplitude of the CCF in the case of FRET can be calculated with Equation 2.33:

CCF

$$N_{CC,app} = \frac{(1-E)(1+E)N_{gr}}{\left(N_g + (1-E)N_{gr}\right)\left(N_r + (1+E)N_{gr}\right)}$$ Equation 2.49

While the error on the total concentrations (as determined from the autocorrelation curve amplitudes) stays below 10% even for 50% FRET efficiency, the error on the concentration of complex [GR] can increase up to 50%, so a good knowledge of possible FRET is crucial for quantitative FCCS analysis. FRET can be quantified by an ensemble fluorescence lifetime measurement by monitoring the decrease of the lifetime of the donor in the presence of the acceptor. However, ensemble results may not be easily translatable to the single protein level as measured in FCCS. For example, RFPs are known to have transient but also long lived dark states, and have been reported to take longer to mature. Reversibly, the effect of FRET on the CCF amplitude can be exploited to quantify the transfer efficiency (Muller et al., 2005).

2.3.4.4. Fast dark states

If one of the fluorochromes in a complex enters a short-lived non-fluorescent dark state, a single-colored particle is temporarily formed. However, this does not reduce the experimental cross-correlation amplitude, as long as the dark state is significantly faster than the diffusion time scale. This can be understood quite easily, if a cross-correlation calculation is made for a discrete dataset:

$$G(\tau) = \frac{1}{N}\sum_{i=1}^{N}\frac{I_{r,i,t}\cdot I_{g,i,t+\tau}}{< I_r >< I_g >}$$ Equation 2.50

If signal I_g is constant in time and signal I_r has transient dark states, their normalized cross-correlation function will not change due to the dark states, since dark states decrease the average $\langle I_r \rangle$ as much as they decrease the correlation term $I_{r,i,t} \cdot I_{g,i,t+\tau}$. In terms of the real cross-correlation amplitude (Equation 2.33), dark states have the same effect on N_{gr} as on $N_{g,total}$ or $N_{r,total}$ so disappear from the equation.

2.3.5. Higher stoichiometries

It is possible to determine the stoichiometry of the protein complex of interest, provided a well characterized FCCS method is used:

- Fluorochromes with dark subpopulations (such as RFPs) should be avoided, because they will complicate the analysis too much
- FRET and cross-talk are better avoided, so the combination of a far-red and a green fluorochrome is best
- Incomplete volume overlap is also best avoided, so it's better not to use lasers that are coupled into one fiber if one wants to use 2-color excitation. If the fluorochrome pair allows it, one can also use 2-photon excitation.

A good (and only) example of stoichiometry determination with FCCS is the *in vitro* and intracellular study of the binding degree, binding affinity and concentrations of Alexa633-labeled Calmodulin (CaM) and Ca^{2+}/CaM-dependent protein kinase II labeled with eGFP under different experimental conditions (Kim et al., 2004; Kim et al., 2005). In general, when considering the following binding reaction:

$$mG + nR \leftrightarrow G_m R_n \qquad \text{Equation 2.51}$$

we can adapt the amplitude of the ACFs (Equation 2.6) and CCF (Equation 2.32) in the presence a higher order stoichiometry, neglecting cross-talk and FRET for reasons of simplicity (Bacia et al., 2002; Kim et al., 2005; Foldes-Papp, 2005):

Green ACF $$G_{green}(0) = 1 + \frac{B_g^2 N_g + (mB_g)^2 N_{gr}}{(B_g N_g + mB_g N_{gr})^2} = \frac{N_g + m^2 N_{gr}}{(N_g + mN_{gr})^2} \qquad \text{Equation 2.52}$$

Red ACF $$G_{red}(0) = 1 + \frac{B_r^2 N_r + (nB_r)^2 N_{gr}}{(B_r N_r + nB_r N_{gr})^2} = \frac{N_r + n^2 N_{gr}}{(N_r + nN_{gr})^2} \qquad \text{Equation 2.53}$$

CCF

$$G_{CC}(0) = 1 + \frac{nmB_gB_rN_{gr}}{(B_gN_g + nB_gN_{gr})(B_rN_r + mB_rN_{gr})} = 1 + \frac{nmN_{gr}}{(N_g + nN_{gr})(N_r + mN_{gr})}$$ Equation 2.54

In excess of red protein, we can calculate:

$$\frac{G_{CC}(0)}{G_{red}(0)} = \frac{\frac{mnN_{gr}}{(0 + mN_{gr})(N_r + nN_{gr})}}{\frac{N_r + n^2N_{gr}}{(N_r + nN_{gr})^2}} = \frac{nN_r + n^2N_{gr}}{N_r + n^2N_{gr}} \xrightarrow{[N_r] \gg [N_{gr}]} = n$$ Equation 2.55

Which means the amplitude of the CCF, relative to the red ACF, will give the binding stoichiometry. In other words, the amplitude of the CCF will be higher than the red ACF amplitude.

On the other hand, in excess of green protein:

$$\frac{G_{CC}(0)}{G_{green}(0)} = \frac{\frac{mnN_{gr}}{(N_g + mN_{gr})(0 + nN_{gr})}}{\frac{N_g + m^2N_{gr}}{(N_g + mN_{gr})^2}} = \frac{mN_g + m^2N_{gr}}{N_g + m^2N_{gr}} \xrightarrow{[N_g] \gg [N_{gr}]} = m$$ Equation 2.56

So, in principle one can determine the binding stoichiometries with FCCS. If the amplitude of the CCF rises above either of the ACFs at excess of one protein, the stoichiometry is always different from 1:1. If considerable cross-talk occurs, at high excess of green protein only the cross-talk will be measured.

2.3.6. Advanced FCCS

One of the primary reasons for performing intracellular FCCS is to determine a protein binding equilibrium constant, free of any artefacts. As discussed, several parameters have to be considered to analyse FCCS measurements correctly, such as the mobility and photobleaching, affinity, volume overlap, cross-talk, quenching or FRET, stoichiometry and finally fluorochrome dark states. It is virtually impossible to control or manipulate all these

parameters. Yet still, great advances in the FCCS methodology have been made to make measurements more transparent.

2.3.6.1. Mobility and stoichiometry

Methods that extend the dynamic time scale of standard point FCCS measurements have been developed, in analogy to FCS (section 2.2.7.3). Not surprisingly thus, spatial image cross-correlation spectroscopy (ICCS) (Wiseman et al., 2000) and the raster image cross-correlation spectroscopy (RICCS) (Digman et al., 2009b) methods were recently introduced to allow probing the interactions in slower complexes. Digman et al. introduced a new method, the Number&Brightness (N&B) analysis, by which both the concentration and stoichiometry of intracellular complexes, irrespective of their mobility, can be determined (Digman et al., 2008). The theory of this technique has also been extended to the two-color case (cross-N&B)(Digman et al., 2009a), to determine the stoichiometry of double labeled complexes. Stoichiometry can also be determined with dual-color PCH (Chen et al., 2005b).

Importantly, non-idealities that complicate the analysis of standard FCS/FCCS measurements will however also have an influence on these novel methodologies, but they do present a powerful addition to the group of protein-protein interaction techniques. Finally, FCCS has also been extended beyond the two-color scheme. Hwang et al. reported a 3-way SW-FCCS method to investigate the interactions between 3 proteins at te same time and Burkhardt et al. even performed a 4-way FCCS experiment by using a grating based detection scheme (Burkhardt et al., 2005; Hwang et al., 2006).

2.3.6.2. Affinity and cross-talk

As for FCS, FCCS has a lower an upper limit to which interactions can be measured (see section 2.2.7). The same methods that can be used to improve the FCS concentration range can be applied for FCCS. Exchanging CW excitation with pulsed excitation and TCSPC based detection greatly increases the precision of FCCS measurements. Furthermore, as we discussed in section 2.3.4.2.2, TCSPC can completely avoid cross-talk by measuring in a PIE fashion. Like this, even weaker affinity binding reactions could be quantified with FCCS.

2.3.6.3. Better fluorescent protein pairs

All photophysical aspects that affect the fluorescence emission of fluorophores will have an effect in an FCCS experiment, which makes FCCS experiment considerably more difficult to interpret than an FCS experiment. As we will discuss in 0, photobleaching, long-lived dark

states and maturation of fluorochromes can have an effect on the observed cross-correlation between two proteins. In general, for quantitative FCCS one best uses a fluorochrome pair with minimal dark state formation, minimal FRET, minimal quenching and minimal cross-talk. The continued development of new fluorescent proteins will eventually lead to a genuine superior fluorescent pair for FCCS.

Part III – Results

Chapter 3. Exploring fluorescence correlation spectroscopy *in vitro*

The challenge of the project was to perform quantitative measurements on proteins inside living human cells. A first goal is to investigate the performance of the FCS technique, originally designed to measure protein concentrations and dynamics in aqueous environment, for quantitative measurements inside living cells. We will do this by exploring experimental conditions in a controlled *in vitro* environment. Possible methodological artefacts, that would be left unnoticed when measuring directly inside living cells, can in this way be identified, controlled for or even eliminated.

Chapter 4. Ultrafast chromatin binding kinetics of LEDGF/p75

In the fourth chapter we focus specifically on chromatin interactions of transcriptional co-activator LEDGF/p75 in living cell nuclei The goal is to characterize the known chromatin binding properties of LEDGF/p75 for the first time in real-time inside living cells, with FCS. By labelling LEDGF/p75 with a green fluorescent protein, eGFP, we can monitor dynamics of the protein in the cell with FCS. The kinetics of chromatin binding will be characterized by customizing the experimental setup. Finally, the role of a major chromatin interacting domain for targeting HIV-1 integrase to chromatin will be investigated, both with FCS, as with a complementary technique for measuring protein dynamics in the slow time scale.

Chapter 5. Measuring the interaction of HIV-1 integrase and LEDGF/p75 in living cells

In the fifth chapter, the interaction of HIV-1 integrase with its co-factor LEDGF/p75 is studied. The goal is to apply the FCCS technique for probing this interaction for the first time in living cells. Since their intranuclear complex is strongly tethered to chromatin, different strategies are followed to probe the interaction with FCCS: the chromatin binding

domains of LEDGF/p75 will be removed, LEDGF/p75 will be retargeted to the cytoplasm, or the major chromatin binding domain of LEDGF/p75 will be altered. Finally, the FCCS technique will be used in a quantitative manner to determine and compare protein-protein interaction affinities. We aim at providing insight into the higher order structure of the complex, the affinity the two protein have for each other and try to relate this to the function for HIV replication.

Chapter 6. FCCS for measuring protein-protein interactions quantitatively in cells

Next, we evaluate the FCCS technique for measuring protein-protein interactions quantitatively in living cells. By combining FCCS measurements on a control system, a genetic fusion construct of two fluorescent proteins, with complementary measurements performed by another technique for quantifying protein-protein interactions, we gain more insight in the biophysical fundamentals of FCCS. Finally, the experimental setup and the fluorochromes used for FCCS are critically evaluated and limitations and possible optimizations are discussed.

Chapter 7. Dark states in monomeric red fluorescent proteins

Finally, we characterize monomeric red fluorescent protein photophysics with FCS, in combination with other spectroscopical methods. We provide structural-functional insight in this matter by linking our observations with crystallographic information.

Chapter 3. Exploring fluorescence correlation spectroscopy *in vitro*

3.1. Introduction

In Chapter 2 we have given an overview of fluorescence correlation spectroscopy (FCS) and have outlined the general biophysical parameters and models that allow to describe and understand a biological system. However, to apply FCS without introducing systematic or random errors, one needs to verify if the autocorrelation function measured under the experimental conditions indeed provides correct and unbiased information. Investigations of the performance of the FCS technique for measuring concentration and diffusion are reported in literature (Hess and Webb, 2002; Enderlein et al., 2004; Enderlein et al., 2005), but it still remains important to test the performance of our specific microscopic setup. In this chapter we test the experimental outcome of an FCS experiment by varying all possible experimental parameters in a controlled *in vitro* environment. We discuss which factors can influence an ACF, independent of the fluorochrome dynamics and concentration. When quantitative intracellular FCS is performed, each of these factors need careful consideration. Successively we investigate the effect of the refractive index, cover glass thickness, temperature, pinhole, viscosity, detector afterpulsing and optical saturation on the ACF.

3.2. Materials and Methods

3.2.1. Buffers

For the refractive index measurements, we made a dilution series of guanidine hydrochloride (GdHCl) in ultra-pure water and measured the refractive index by refractometry (Nozaki, 1972)(Table 3.1). For the viscosity measurements, a dilution series of glycerol in ultrapure water or phosphate buffered saline was used (Table 3.2). Occasionally, phosphate buffered saline pH 7.4 (PBS) was also used.

Table 3.1 **Refractive index of GdHCl**

Concentration (M)	n (calculated)	n (measured)
0.575	1.34	1.340
1.158	1.35	1.350
1.747	1.36	1.360
2.342	1.37	1.370
2.943	1.38	1.380
3.548	1.39	1.390
4.158	1.40	1.401
4.772	1.41	-
5.390	1.42	-

Table 3.2 **Viscosity of glycerol**

Glycerol (w/w)	η at 20 °C (cP-mPa·s)
0	1.005
40	3.72
75	35.5
93	367

3.2.2. Fluorochromes

Rhodamine 6G (Laser grade, Acros Organics BVBA, Geel, Belgium) and Alexa 488 succinimidyl ester, named 'Alexa 488' from here on (N.V. Invitrogen, Merelbeke, Belgium) were dissolved in dry DMSO at 10 mg/mL and stored at -20 °C. For experiments, the fluorochromes were diluted to 1-50 nM in ultrapure water or buffer. His-tagged enhanced green fluorescent protein was obtained by bacterial expression from pEXP5-eGFP (kind gift from Dr. Rob Lavigne, Afdeling Gentechnologie, Katholieke Universiteit Leuven) and His-tagged monomeric red fluorescent protein was obtained by bacterial expression from pRSET-mRFP1 (kind gift of Dr. Roger Y. Tsien (HHMI-UCSD, La Jolla, CA). After transformation of the plasmid in E. coli BL21 cells, the cells were grown at 37 °C to an optical density of 0.6 after which protein overexpression was induced during 3 h with 1 mM IPTG. After sonication of the culture, the proteins were purified using gravity flow Ni^{2+}-affinity chromatography (Protino Ni-TED, Macherey-Nagel GmbH & Co. KG, Düren, Germany). The protein purity was checked with SDS-PAGE. Fluorescent proteins were dissolved in appropriate buffer.

3.2.3. Cover glasses

Lab-Tek™ Chambered Cover glasses (#1 d=0.13-0.16 mm) with 8 chambers (VWR International, Leuven, Belgium) were used for the cover glass thickness measurements. For all other measurements we used normal #1 cover glasses (Waldemar Knittel Glasbearbeitungs GmbH, Bielefeld, Germany).

3.2.4. FCS measurements

FCS measurements were performed on a commercial ConfoCor2 system (Carl Zeiss, Jena, Germany). The 488-nm line of an Ar-ion laser was used to excite Alexa 488, rhodamine 6G and eGFP. The 543-nm HeNe laser was used to excite rhodamine 6G and mRFP1. Excitation light was reflected by a 488/543 dichroic mirror and focused in the sample through a C-Apochromat 40x-NA1.2-W objective (Carl Zeiss). Excitation intensity was regularly measured at the back aperture of the objective with a power meter (Orion PD-300, BFi OPTiLas, Alpen aan den Rijn, The Netherlands), to ensure proper laser alignment in the fiber coupling unit (Figure 3.1). Also, because experimental parameters of fluorochromes, especially fluorescent proteins (section 7.3.1.1), are dependent on the excitation intensity, this ensures reproducibility of the results.

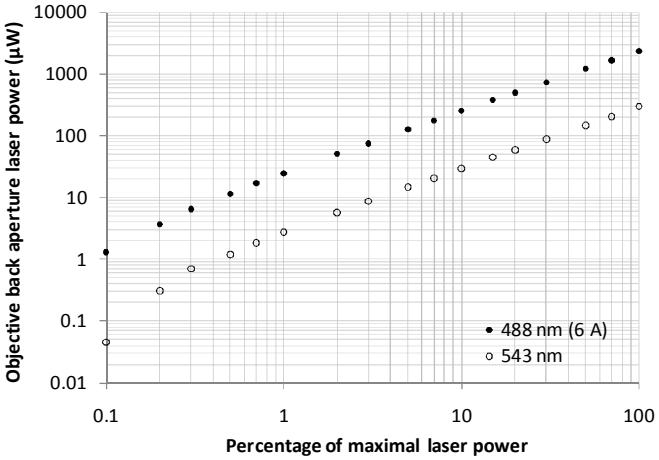

Figure 3.1 **Laserpower versus AOTF (acousto-optical tunable filter)** – Experimentally measured laser power of 488-nm Ar-ion and 543-nm HeNe at the back aperture of the objective.

Laser intensity is controlled with an acousto-optical tunable filter (AOTF). Fluorescence in the green channel passed through a 505-530 bandpass filter and was focused on an avalanche photodiode (APD) through a 70-µm pinhole. Fluorescence in the red channel passed through a 600-650 bandpass filter and was focused on another APD through a 78-µm pinhole. For the different experimental conditions 10 measurements of 20 seconds were performed and the average autocorrelation curve was calculated. Measurements were performed at room temperature (21 °C) unless stated otherwise. Measurements were analysed in Igor Pro 5 (Wavemetrics Inc., Lake Oswego, OR, USA).

3.3. Results

3.3.1. Refractive index mismatch

When focused excitation light from the objective traverses the cover glass it is refracted into the sample solution. Under normal circumstances, when the refractive index of the immersion solution is the same as that of the sample solution, the objective focuses the light optimally. For a water-immersion objective the refractive index should thus be n=1.33. Although the refractive index of most buffers does not differ much from that, the intracellular environment can have a n=1.35-1.40 (Brunsting and Mullaney, 1974; Bereiter-Hahn et al., 1979; Lanni et al., 1985; Curl et al., 2005).

Enderlein et al. reported alarming numerical quantifications of the effect such refractive index mismatches would have on the obtained parameters from FCS measurements (Enderlein et al., 2005). To test this experimentally, we performed an experiment where we incubate HeLa cells that had been cultured on a cover glass with PBS containing a nM-concentration of eGFP. We then measured the ACF in this medium, well above a cell and at certain positions with respect to the cell. Because the concentration of the probe logically is independent of the location, the effect of a cell in the optical path on the experimental parameters can be directly observed (Figure 3.2).

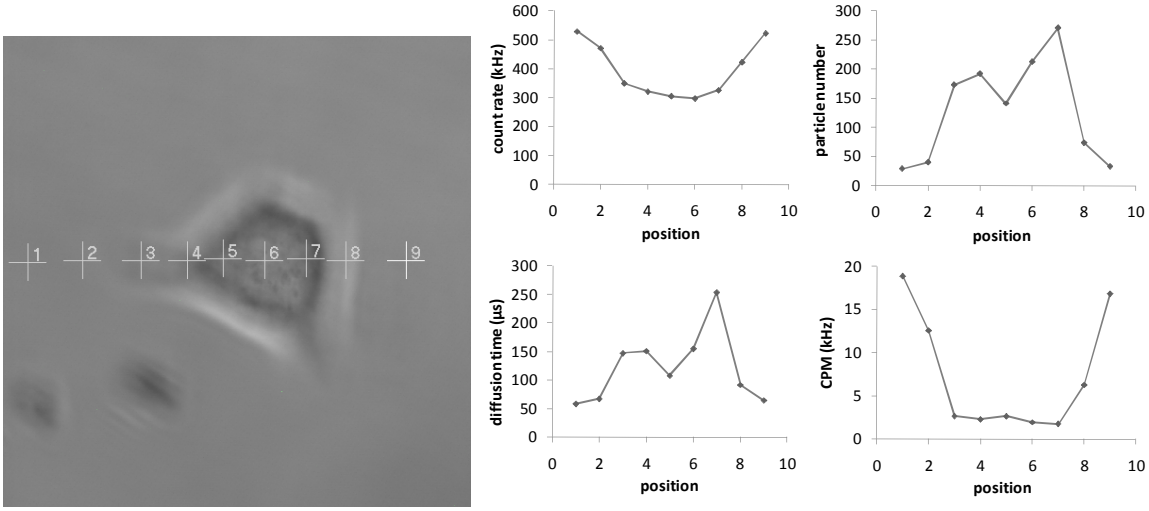

Figure 3.2 **FCS of extracellular eGFP in the presence of cellular refraction** – (left) LSM image of a HeLa cell taken 15 µm above the cover glass, which makes the cell appear out-of-focus. Crosses indicate the FCS measurement positions. (right) FCS measurement of a 50-nM solution of eGFP in PBS at 15 µm above the cover glass. The underlying cell had a maximum height of 11.5 µm.

The experimental count rate depended strongly on the position with respect to the cell, suggesting a defocussing of the excitation/emission light by the cell. The increasing particle number and diffusion time also confirmed an enlarging focal spot and the experimental CPM (Equation 2.7), which is an excellent sensor for the quality of the focal spot, decreased considerably. We thus concluded that the presence of the cell influences solution FCS measurements of eGFP considerably. Furthermore, because this effect was so strong, we suspected that also intracellular parameters such as the diffusion coefficient and the concentration were heavily biased by the differential refractive index.

Next, to investigate this effect more quantitatively, we performed an FCS experiment on a 20-nM solution of rhodamine 6G, in a series of buffers with different refractive indices and at different heights above the cover glass. It is immediately obvious that ACFs measured at 10 μm above the cover glass suffered from a relatively small, yet significant effect of the refractive index, while the measurements at 100 μm were drastically influenced even upon a small refractive index mismatch (Figure 3.3 A and B). The experimental ACFs were fit with Equation 2.17, to check the effect of the refractive index on the individual parameters. As can be seen in Figure 3.3 C, the experimental count rate decreased, even at 5 μm from the cover glass. This is partially due to a quenching effect of the guanidine hydrochloride at high concentration. Figure 3.3 D and E show that at 5-10 μm distance from the cover glass, which is about the height at which a typical intracellular measurement is performed, the particle number and diffusion times are not influenced much by the refractive index. For example, at 10 μm in a refractive index similar to the intracellular environment (n=1.38) the particle number was decreased about 13% from that at n=1.33, in accordance with reported suggested values (Enderlein et al., 2004). Further from the cover glass, at 100 μm, this amplitude was decreased by almost 83%. A larger n also had an effect on the viscosity of the solution (about a factor of 1.4 for n=1.42), explaining the increase of the diffusion time in water. To relate the observations with actual defocussing of the objective, we calculated the size of the confocal volume from the experimental particle numbers (Figure 3.3 D inset). The confocal volume increased from 0.53 fL up to 6 fL, when measuring far from the cover glass in a buffer with a large refractive index mismatch. Because a more diffuse focal spot means a lower light/area, a decrease in counts-per-molecule (CPM) was also observed. Of importance, this decrease in CPM could not be related only to an effect of GdHCl on the optical properties of rhodamine 6G, since close to the cover glass the brightness was much less affected by GdHCl. In conclusion, for *in vitro* measurements it is always better to stay

close to the cover glass, because potential artefacts due to refractive index mismatches are almost completely avoided. Intracellular FCS measurements do provide an adequate (within 10% error at 5 µm height) estimation of the concentration and dynamics given a particular *in vitro* calibration of the setup.

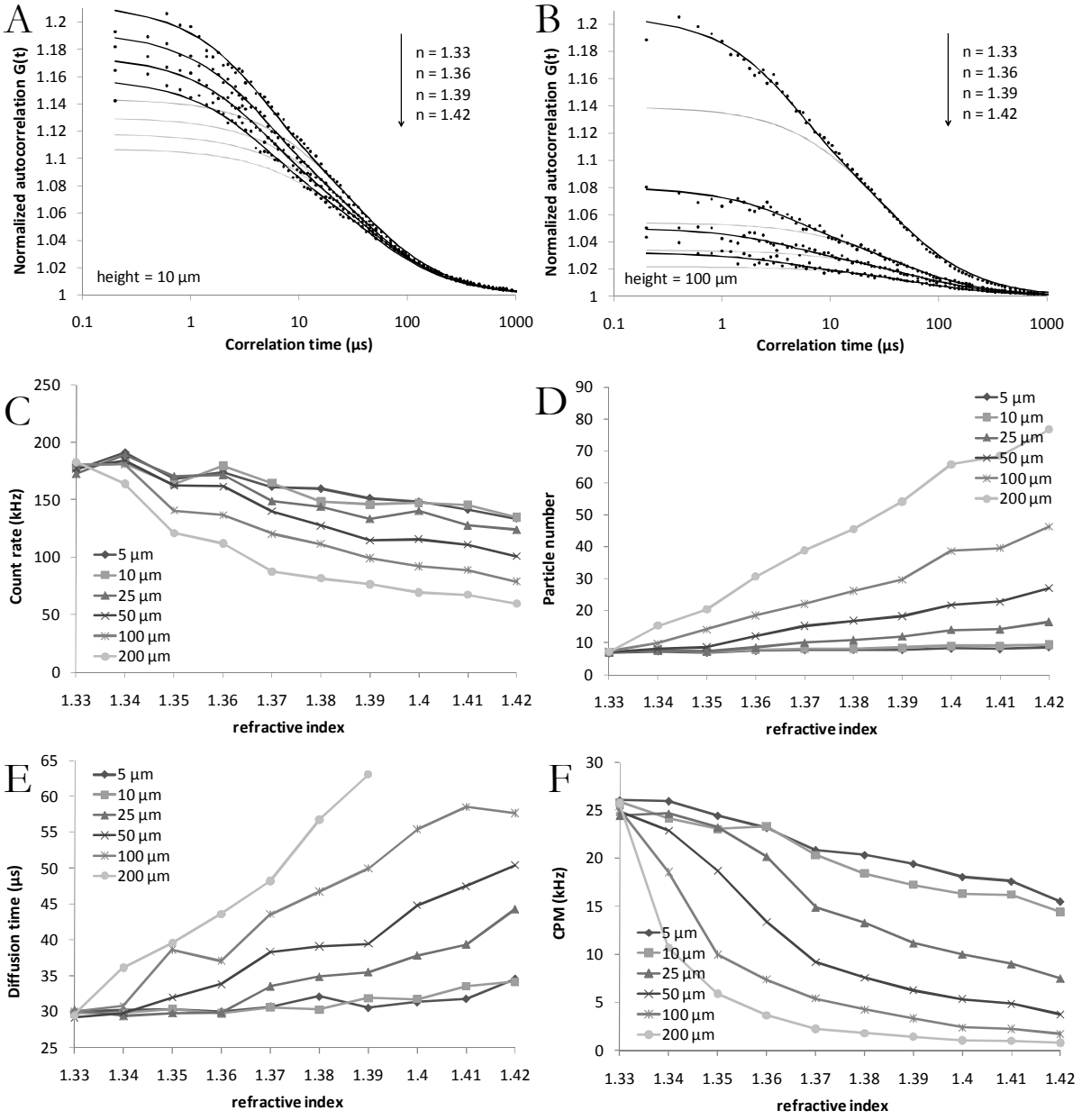

Figure 3.3 **Effect of refractive index on FCS of rhodamine 6G** – ACFs were measured at different heights above the cover glass on a solution of 20 nM rhodamine 6G in water. Panels (A) and (B) exemplify the ACFs measured at resp. 10 µm and 100 µm in buffer with different refractive indices. The gray fit line is the concentration amplitude of the ACFs. Different parameters obtained after fitting: (C) average count rate, (D) particle number and (inset) confocal volume, (E) diffusion time and (F) counts-per-molecule.

3.3.2. Cover glass thickness

An FCS objective has a correction collar to adjust the focussing of the objective to the actual thickness of the cover glass. We verified the objective collar-cover glass thickness mismatch experimentally by performing FCS measurements on Alexa 488 in water at various positions of the collar, each time at 10 μm from the cover glass. As a cover glass we used the cuvettes for intracellular measurements. In principle, these '#1' 0.13-0.16 mm cover glasses should have an average thickness of 0.145 mm.

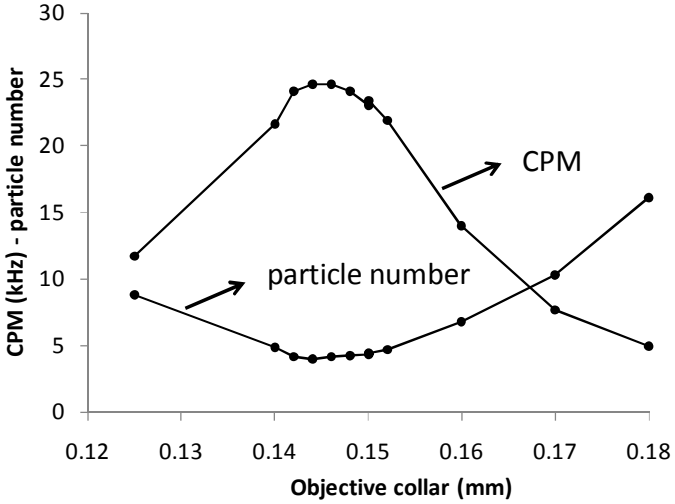

Figure 3.4 **Effect of objective collar setting on FCS of Alexa 488** – ACFs were measured at 10 μm above the #1 cover glass on a 20-nM solution of Alexa 488 in water. CPM: counts-per-molecule.

As can be seen in Figure 3.4, both the experimental CPM as the particle number clearly depends on the setting of the correction collar. An optimal focussing (lowest particle number and highest brightness-per-molecule) was found for a collar setting of 0.145 mm, which is exactly the average thickness. The particle number is quite constant in a range of +/- 5 μm of cover glass thickness.

We thus conclude that an FCS setup is better calibrated with identical cover glasss to the ones used for intracellular measurements. Doing the calibration measurement in an empty chamber of the same cover glass containing the cells is even recommended, although a large variation on the actual thicknesses was not observed.

3.3.3. Temperature

When FCS measurements are to be performed in living cells, it is sometimes necessary to work at 37 °C. The effect of the temperature on the experimental diffusion coefficient will first be briefly discussed, after which the practical implementation of 37 °C incubation and the effect on the microscope optics will be discussed.

The diffusion coefficient is directly related to the temperature, but inversely related to the viscosity of the medium (Equation 2.9). For water, the dynamic viscosity is empirically approximated by (Vogel, 1921):

$$\eta(T) = A \times 10^{\frac{B}{T-C}} \qquad \text{Equation 3.1}$$

where A=2.414×10^{-5} Pa·s, B = 247.8 K and C = 140 K.

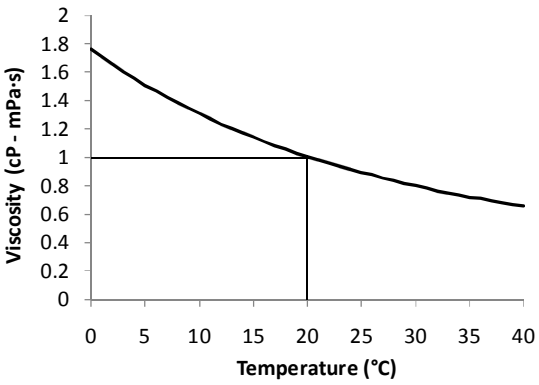

Figure 3.5 **Temperature dependence of the viscosity of water** (Vogel, 1921).

If the temperature is increased from 21 °C to 37 °C, the diffusion coefficient of a fluorochrome *in vitro* will thus increase by 49.4%. In living cells however, the diffusion coefficient of a protein might not have the same temperature dependence as predicted by the Stokes-Einstein law, since some parts of the cellular metabolism might be more dependent on temperature than others. Also, at 37 °C living cells exhibit a higher mobility than at room temperature, which in some cases might even complicate FCS measurements.

Next, for stable and uniform 37 °C incubation of the sample, both the cuvette containing the cells as well as the objective need to be warmed. High-NA objectives act as a strong heat-sink, because they make physical contact with the sample through the immersion solution. Since some cellular reactions are strongly temperature dependent, an absolute, homogenous and reproducible temperature distribution needs to be assured in all positions of the cuvette,

which in our case were Lab-Tek™ Chambered Cover Glasses. Because commercial cuvette heaters did not provide this, a custom heater that provided equal heating of each chamber was constructed. A commercial objective heater was also used (Bioptechs Inc., Butler, PA, USA). We measured the ACF of a 15-nM Alexa 488 solution with the 488-nm line of the Ar-laser, or a 10-nM rhodamine 6G solution with the 543-nm HeNe laser at either room temperature (RT) or 37 °C, and varied the cover glass-thickness correction collar of the objective. When the temperature of the sample and objective was raised to 37 °C, a 30-minute incubation period was applied after the set temperatures had been reached, to allow complete equilibration of the system.

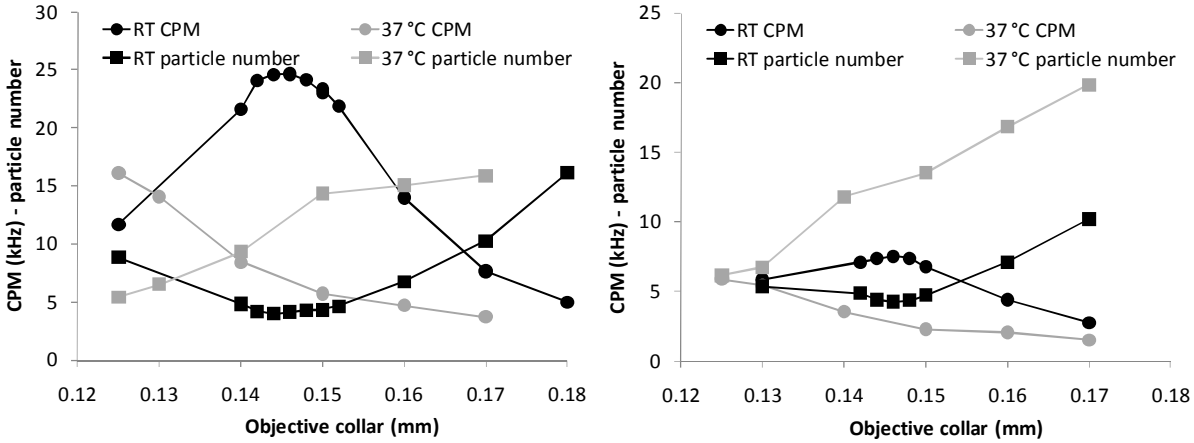

Figure 3.6 **Effect of temperature on the objective optical properties** – Apparent particle number and counts-per-molecule at different objective collar settings for a (left) 15-nm Alexa 488 or (right) 10-nM rhodamine 6G solution in water at room temperature or 37 °C.

As can be seen in Figure 3.5, a rather large effect on the focussing of the excitation beam was observed for Alexa 488 at 37 °C, since the counts-per-molecule dropped by a factor of more than 3, for the standard setting of the collar (0.145 mm). For rhodamine 6G we observed a similar defocussing of the objective. By turning the correction collar to smaller values, we could partially recover the focussing, reaching up to 70% of apparent brightness for Alexa 488 and 80% for rhodamine 6G. Taking the diffusion coefficients of Alexa 488 and rhodamine 6G (Table 2.1) and the temperature into account, we calculated the size of the confocal volume (Table 3.3). Both at 488-nm excitation as 543-nm excitation the correction collar could not achieve the optimal focussing achieved at room temperature, although at 488-nm excitation the difference was only small.

Table 3.3 **Size of the confocal volume at room temperature and 37 °C under optimal focussing conditions** – τ_{diff} = diffusion time, $\omega_{1,2}$ = radial and axial radii of the focal volume, V = confocal volume.

		τ_{diff} (µs)	ω_1 (nm)	ω_2 (nm)	V (fL)
alexa fluor 488	RT	22.4	197	1160	0.29
	37 °C	16.8	209	1170	0.32
rhodamine 6G	RT	32.8	236	1110	0.39
	37 °C	26.7	260	1250	0.53

Summarized, it is possible to measure at varying temperatures but the optical properties of the objective will be influenced. The objective need to be carefully characterized when working at 37 °C, especially if results at RT and 37 °C are to be compared. Also, the volume overlap for FCCS measurements needs careful characterization when working at 37 °C.

3.3.4. APD afterpulsing

In avalanche photodiodes (APD, SPAD), the detectors mostly used for FCS measurements, an emission photon is detected and converted into an electrical pulse. However, nanoseconds to microsecond after the initial pulse is generated, these detectors also generate a so-called 'after-pulse'. When the signal-to-noise ratio is low, i.e. the propability of detecting a 'real' photon is low due to a low concentration or brightness of the fluorochrome, this afterpulsing appears in the autocorrelation curve at a (sub-)microsecond time scale. We experimentally verified APD afterpulsing on the Zeiss ConfoCor2 microscope by measuring a low concentration of rhodamine 6G at increasing laser powers. As the concentration of the probe remains constant, the photophysical phenomena in the ACF can be verified. As can be seen in Figure 3.7, at low count rates the afterpulsing is clearly detectable as a sharp rise of the ACF in the submicrosecond time scale and obscures the contribution of µs-time scale triplet conversion to the ACF. At higher count rates (shown by the arrow), the ACF becomes gradually less and less distorted so that even submicrosecond ACF datapoints are unbiased by the afterpulsing phenomenon.

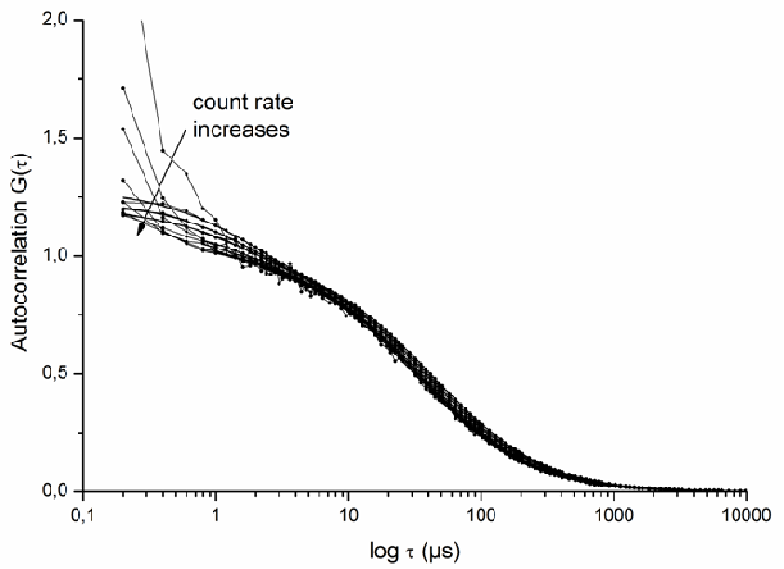

Figure 3.7 **Influence of detector afterpulsing on the experimental ACF** – A low-concentration rhodamine 6G solution was measured with FCS at rising 543-nm HeNe laser powers, shown by the arrow.

It is thus possible to minimize artefacts due to APD afterpulsing, by measuring at count rates where the afterpulsing does not contribute significantly to the signal. We have done this when studying the dark states of mRFPs (Chapter 7). Complete avoidance of afterpulsing from the ACF is possible by measuring in cross-correlation mode. Here, the emission of the fluorochrome is split 50/50, detected on two APDs and cross-correlated. At low count rates, that would give afterpulsing in autocorrelation mode (10 kHz), a photon is detected only once every 100 µs, so the chance that the two detectors in cross-correlation mode register a photon (and an afterpulse) at the same time is very low. In other words, afterpulsing does not cross-correlate between detectors and hence does not influence the ACF when measuring in cross-correlation mode.

3.3.5. Optical saturation

At low laser powers the emission intensity is linearly proportional to the excitation intensity. At higher laser powers on the other hand, the excited state population starts to saturate (Enderlein et al., 2005). We studied the excited state saturation of rhodamine 6G with FCS. In the previous section we measured a low concentration of rhodamine 6G at increasing laser powers. In Figure 3.8 the parameters obtained after fitting of these measurements are given. The counts-per-molecule levels off at high laser powers, clearly indicating optical saturation (Figure 3.8 A). Importantly, the diffusion time of rhodamine 6G steadily increased at increasing laser powers (Figure 3.8 B). Indeed, due to the Gaussian shape of the excitation

profile, an apparent increase of the confocal volume is observed under optical saturation conditions, explaining the increased diffusion time (Gregor et al., 2005). Since more molecules are observed in a larger volume, this also explains why the count rate seems to level off less than the CPM (Figure 3.8 C). Finally, as the excited state population increases with the laser power, so does the triplet state population (Figure 3.8 D).

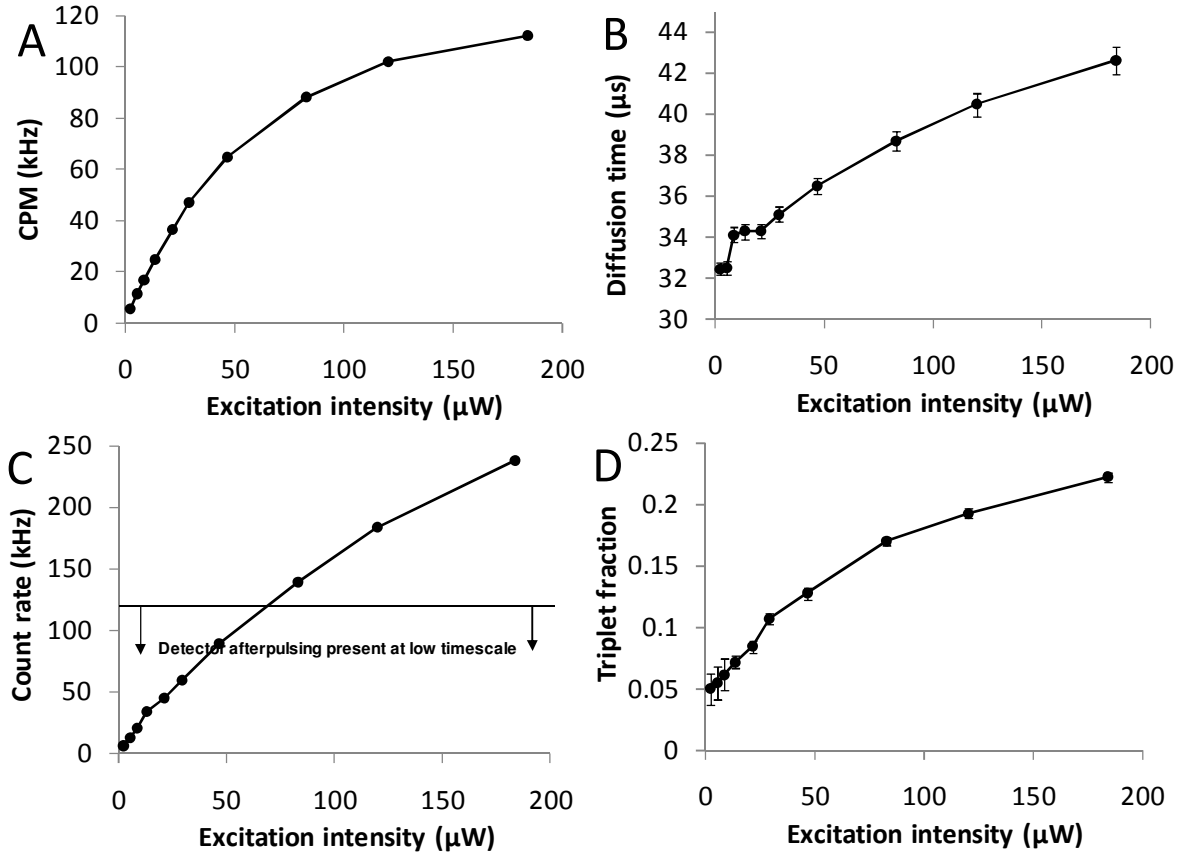

Figure 3.8 **The effect of optical saturation on the experimental ACF of rhodamine 6G** – A low-concentration of rhodamine 6G was measured with FCS at increasing 543-nM HeNe laser powers and the ACFs were analysed.

In conclusion, optical saturation leads to an apparent increase in the laser focus. To avoid this it is better to measure ACFs always at low excitation intensity. Importantly, optical saturation depends not only on the laser intensity but also on the fluorochrome that is being used. Calibrating the microscope with one probe at a certain laser intensity might thus result in a different effective confocal volume for another probe, if it exhibits different optical saturation characteristics.

3.3.6. Volume overlap

As we have discussed in section 2.3.4, different wavelengths of excitation are focused by the objective in diffraction limited excitation spots of unequal size. Since the coupling of the laser beams into the objective back aperture could not be modified, we wished to develop a simple, robust and reproducible procedure to calibrate the size of the confocal volumes. In each emission path there is a pinhole, to allow independent tuning of the size of the confocal volumes. We verified the effect of the pinhole diameter of the two detection channels on the experimental parameters from the ACF (Figure 3.9).

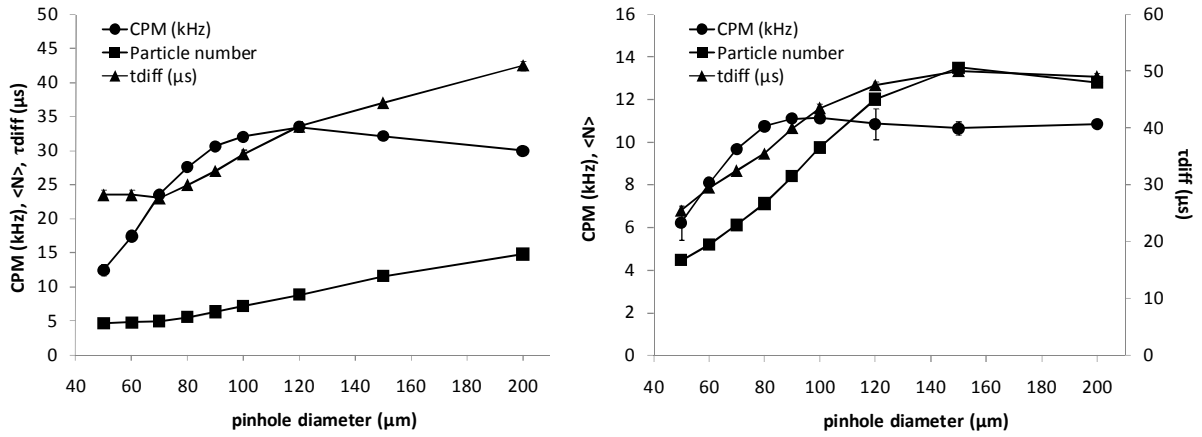

Figure 3.9 **The effect of the pinhole diameter on the experimental ACF** – (left) A low concentration of Alexa 488 was measured with 488-nm excitation and (right) a low concentration of rhodamine 6G was measured at 543 nm excitation. The experimental counts-per-molecule, average particle number and diffusion time are depicted.

As expected, when increasing the pinhole diameter a larger confocal volume is 'seen', translated in a larger diffusion time and particle number. The counts-per-molecule however, reaches a maximum at ~90 µm for 488-nm excitation of Alexa 488 and of about 80 µm for 543-nm excitation of rhodamine 6G. Importantly, when the pinholes are set at their defaults values (as indicated in the manual of the microscope), 70 µm for cyan excitation and 78 µm for green excitation, and equilibrium between high CPM and small focus is guaranteed. Thus, by tuning the individual pinhole diameters one can have control over the effective size of the two confocal volumes. Incomplete volume overlap (section 2.3.4) can in this way be partially corrected for.

Next, we wanted to characterize the confocal volumes created by the 488-nm line of the Ar-ion laser and by the 543-nm HeNe laser. First, we calibrated both pinholes in XYZ direction at the respective excitation wavelengths, 488 nm for the first pinhole, 543 nm for the second pinhole. Then we measured rhodamine 6G with 543-nm excitation and split the emission 50/50 in the two channels (600-650 nm, 78-µm pinhole). The particle number from the ACF in the two channels was the same, which proves the 543-nm laser focuses at the same spot as does the 488-nm laser. This confirmed that both pinholes image the same spot in solution. Then we tested if rhodamine 6G could be used to calibrate the two confocal volumes. As can be seen in Figure 3.10, the 488-nm excitation intensity needed to obtain reliable parameters from the ACF measured in the green detection channel, about 7% of the maximal intensity, induces strong optical saturation, as judged from the CPM. In other words, it is not possible to obtain a reliable estimation of the absolute parameters (S, N, τ_{diff}) at an excitation intensity that does not introduce optical artifacts. The 543-nm excitation intensity needed for reliable parameters estimation in the red detection channel on the other hand, which was about 15% of the maximal intensity, did not introduce any optical artefacts yet, since the extrapolated values are much alike the ones measured at 15%. In other words an optical artifact free particle number and diffusion time of rhodamine 6G can be measured in the red detection channel by doing a calibration measurement at 15% of the maximal 543-nm power. Only the size of the red confocal volume can thus be determined accurately with rhodamine 6G.

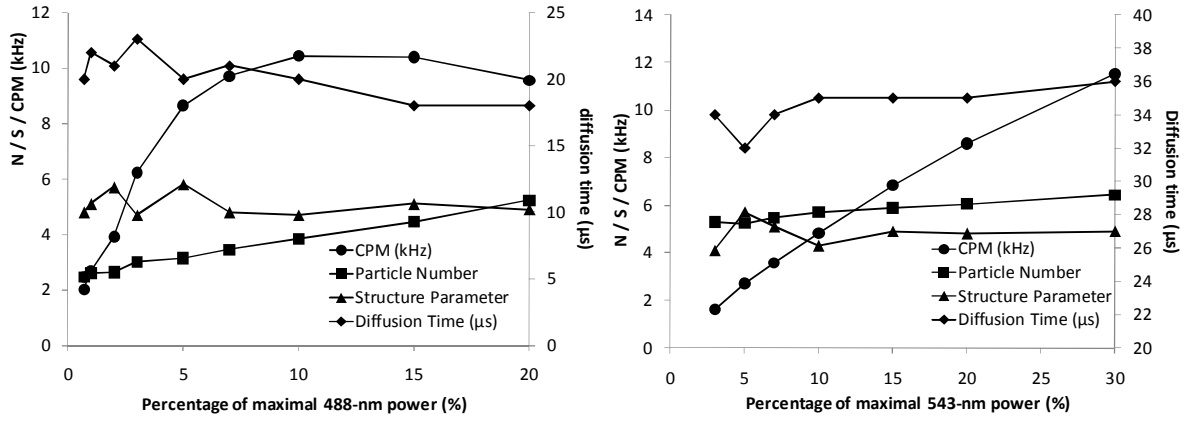

Figure 3.10 **Excitation power needed to obtain reliable parameter estimations from the rhodamine 6G ACF** – (left) 488-nm excitation, a 505-530-nm bandpass emission filter and 70-µm pinhole. (right) 543-nm excitation, a 600/650-nm bandpass emission filter and 78-µm pinhole.

For the green detection channel we then used Alexa 488 instead of rhodamine 6G. Already at 1% of the maximal 488-nm intensity parameters could be estimated with an error of less than 10% due to optical artifacts. At this laser intensity, optical saturation of Alexa 488 is thus still limited, as also judged from the CPM.

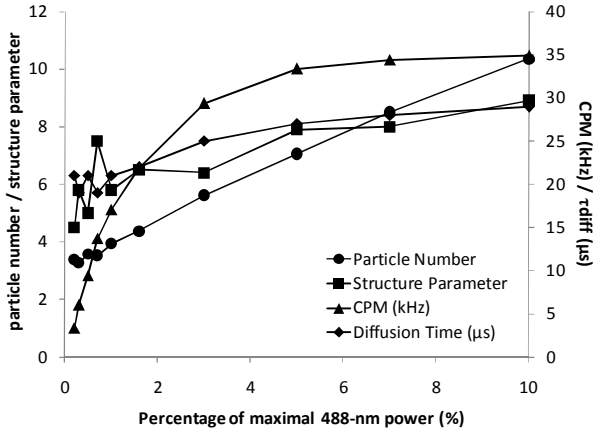

Figure 3.11 **Excitation power needed to obtain reliable parameter estimations from the Alexa 488 ACF** – 488-nm excitation, a 505-530-nm bandpass emission filter and 70-µm pinhole.

Importantly, if the 488-nm/green channel is characterized correctly with Alexa 488 and the 543-nm/red channel with rhodamine 6G, then the ratio of calculated confocal volumes (Figure 3.10 right and Figure 3.11) should be equal to the ratio of particle numbers in each channel as measured for rhodamine 6G alone (Figure 3.10 left and right). Table 3.4 lists the parameters from the fitting. After linear extrapolation to 0 light intensity, we find that the ratio of volumes was 2.08, while that of the particle numbers was 2.10. Importantly, at the laser powers used for calibrating the volumes, 1% of the 488-nm laser power and 15% of the 543-nm laser power, this ratio only differed from 2.08 by 13%, which is within the error of cellular measurements.

Table 3.4 **Comparison of the ratio of rhodamine 6G particle numbers, as measured with 488-nm and 543-nm excitation, with the ratio of calculated confocal volumes, as measured with Alexa 488 and rhodamine 6G** – The values in bold are the values extrapolated to zero excitation intensity.

rhodamine 6G excitation: 488 nm emission: 505-530 nm			rhodamine 6G excitation: 543 nm emission: 600-650 nm			Alexa 488 excitation: 488 nm emission: 505-530 nm		
I_{exc} (%)	N	V (fL)	I_{exc} (%)	N	V (fL)	I_{exc} (%)	N	V (fL)
0.70	2.47		3.00	5.30		0.20	3.39	0.20
1.00	2.60	0.24	5.00	5.23	0.47	0.30	3.26	0.22
2.00	2.64	0.25	7.00	5.46	0.46	0.50	3.59	0.22
3.00	3.00	0.24	10.00	5.71		0.70	3.53	0.28
5.00	3.15	0.24	15.00	5.88	0.46	1.00	3.94	0.25
7.00	3.45	0.21	20.00	6.05	0.45	1.60	4.38	0.31
10.00	3.83	0.19	30.00	6.43	0.48	3.00	5.62	0.36
15.00	4.46	0.18				5.00	7.06	0.51
20.00	5.23	0.17				7.00	8.50	0.54
						10.00	10.37	0.63
	2.45	**0.25**		**5.16**	**0.46**		**3.20**	**0.22**

In conclusion, the pinholes can be fine-tuned to image a smaller or bigger focal spot in solution. In a calibrated system, these pinholes image exactly the same spot in solution. The two confocal volumes created with 488-nm (AOTF 1%) and 543-nm (15%) excitation in a dual color FCCS setup can be determined accurately with Alexa 488 and rhodamine 6G.

3.4. Discussion – Conclusion

Ideally, the autocorrelation function of a certain protein in a cell as measured with FCS is determined only by the concentration and dynamics of the protein, as the biophysical theory we have outlined in Chapter 2 would lead us to assume. In practice, as we have seen in this chapter, a number of experimental parameters also critically determine the ACF. A good characterization of the microscopic setup was therefore important before experiments in cells could really be trusted. We showed that although the intracellular environment, in terms of optical density, is much different from an aqueous environment, the error introduced on experimental parameters is still acceptable.

Next, we showed that the cover glass thickness has a strong effect on the ACF and that the correction collar of the objective is best optimized for the cover glass to be used. Also, when experiments are performed at 37 °C, the performance of the system is still acceptable when the objective collar is re-optimized. In a two-pinhole system such as the Zeiss ConfoCor2 FCS microscope, the individual pinholes can be adjusted to fine-tune the size of the confocal volume in a limited range, which can potentially increase the volume overlap when doing 2-color FCCS measurements. We have shown that at low signal-to-noise detector afterpulsing can negatively influence the ACF at fast time scale and that measurements at high excitation intensities can suffer from optical saturation artefacts. FCS measurement are thus best performed at excitation intensities that are high enough to have a good CPM and low afterpulsing artefacts, but low enough to avoid optical saturation. Finally, we have provided an accurate method to characterize two confocal volume used for 488/543 dual color FCCS measurements.

Chapter 4. Ultrafast chromatin binding kinetics of LEDGF/p75

Part of this chapter will be published in:

Hendrix, J., De Rijck, J., Gijsbers, R., Voet, A., Hotta, J., McNeely, M., Vanstreels, E., Daelemans, D., Hofkens, J., Debyser, Z., and Engelborghs, Y. Dynamic chromatin scanning of transcriptional co-activator LEDGF/p75 is arrested by HIV-1 integrase (submitted for publication).

4.1. Introduction

Gene expression can be regulated by transcriptional cofactors, which fine-tune the interaction of the general transcription machinery with gene-specific transcription factors. Lens epithelium-derived growth factor/p75 (LEDGF/p75) was originally identified as a transcriptional co-activator protecting cells against apoptosis (Ge et al., 1998a). LEDGF/p75 expression is upregulated under thermal or oxidative cellular stress (Sharma et al., 2000) and in turn, promotes transcription of stress-related genes by binding to promoters containing stress responsive (STRE) or heat shock elements (HSE) (Singh et al., 2001). The protein plays an important role in oncogenesis (Ahuja et al., 2000; Daugaard et al., 2007; Huang et al., 2007; Yokoyama and Cleary, 2008), autoimmunity (Ganapathy and Casiano, 2004; Daniels et al., 2005) and integration and replication of the human immunodeficiency virus type 1 (HIV-1) (Van Maele and Debyser, 2005; Engelman and Cherepanov, 2008).

LEDGF/p75 belongs to the family of HDGF-related proteins characterized by a N-terminal PWWP-domain, a conserved DNA/chromatin-binding domain (Stec et al., 2000; Singh et al., 2006). Indeed, eGFP-PWWP does interact with mitotic chromosomes and deletion of this domain from full-length LEDGF/p75 decreases its affinity for DNA/chromatin *in vitro* (Llano et al., 2006b) and in cell culture (Turlure et al., 2006; Llano et al., 2006b). A tripartite

element consisting of two AT-hooks and the nuclear localisation signal (NLS) cooperates with the PWWP-domain for DNA/chromatin interaction *in vitro* (Llano et al., 2006b) and in cell culture (Turlure et al., 2006). The integration of HIV in the genome of the infected host cell is catalyzed by integrase (IN) in close interaction with LEDGF/p75 (Cherepanov et al., 2003). This interaction maps to the conserved integrase binding domain (IBD) at the C-terminus of LEDGF/p75 (Cherepanov et al., 2004; Cherepanov et al., 2005a). Potent knockdown or knockout of LEDGF/p75 drastically hampers HIV replication (Llano et al., 2004b; Vandekerckhove et al., 2006; Zielske and Stevenson, 2006; Marshall et al., 2007; Shun et al., 2007). Multiple roles in the HIV integration process have been attributed to LEDGF/p75: (i) protection of IN against cytoplasmic proteasomal degradation (Llano et al., 2004a), (ii) promotion of the oligomerisation of IN to allow DNA strand transfer (Faure et al., 2005; McKee et al., 2008; Michel et al., 2009), (iii) tethering of the HIV-1 pre-integration complex to the chromatin (Turlure et al., 2006; Llano et al., 2006b) and (iv) targeting of integration into actively transcribed regions of the genome (Ciuffi et al., 2005; Marshall et al., 2007; Shun et al., 2007).

In vivo chromatin binding is classically studied with Fluorescence recovery after photobleaching (FRAP)(Sprague et al., 2004; Phair et al., 2004) and very recently, Fluorescence correlation spectroscopy (FCS) was also applied for this purpose (Michelman-Ribeiro et al., 2009). Despite the fact that FCS analyzes diffusion on a much faster (μs-ms) timescale than FRAP (seconds) and might thus be a better method to study chromatin binding kinetics, it is still arguable that these kinetics might even be faster than quantifyable with standard FCS (Michelman-Ribeiro et al., 2009). In this chapter, we investigate chromatin binding kinetics of LEDGF/p75 with tunable focus FCS to show extraordinary fast chromatin binding kinetics.

4.2. Materials and Methods

4.2.1. Plasmids

For the expression of eGFP in human cells we used a peGFP-C1 construct (Clontech Laboratories, Inc., Mountain View, CA), that allows for easy expression of eGFP labelled genetic fusion constructs downstream of a cytomegalovirus immediate early promoter. The molecular clonings of the constructs for the eukaryotic expression of eGFP-LEDGF/p75, eGFP-LEDGF/p75 K150A, eGFP-Δ325, eGFP-Δ325 D366A and eGFP-IN in HeLa cells

are described elsewhere (Maertens et al., 2003; Maertens et al., 2004; De Rijck et al., 2006). The PWWP mutants were constructed by site-directed mutagenesis of pCP-Nat75 (Maertens et al., 2003). Lysine56 was mutated to aspartic acid using the Kirsch-Joly method (Kirsch and Joly, 1998). First, two 483-bp primers were synthesized by PCR with two primers: 5'-GAGA CTGCTTTTTTAGGACCAGACGATATCTTTCCTTACTCAGAAAATAAGG (sense, mutated positions underlined) and 5´-GTCACTCTCTGAAGGAC (antisense). These primers were then used to synthesize the whole plasmid in an additional PCR reaction, after which template DNA was digested with *Dpn*I. Colonies were screened by restriction and sequence analysis. Arginine74 was mutated to aspartic acid using the same protocol with 5´-GGAAAAGTATGGCAAACCAAATAAAGACAAAGGTTTTAATGAAGGTTTATGG G (sense) and the same antisense primer to form two 433-bp primers. Colonies were screened by sequence analysis. The pCP-Nat75 K56D-R74D construct was obtained by mutagenesis of the pCP-Nat75 R74D construct with the K56D primer set.

Protein purification was performed as for wild-type LEDGF/p75 (Maertens et al., 2003). For imaging inside live cells, the mutant LEDGF/p75 genes were PCR amplified with 5'-CACGAGATCTGACTCGCGATTTCAAACCTGGAGACC (sense) and 5'- CCGCGAA TTCTAGTTATCTAGTGTAGAATCCTTC (antisense) and cloned into peGFP-C3 via *Bgl*II and *Eco*RI restriction sites. Clones were verified with restriction and sequence analysis. A construct for the eukaryotic expression of mRFP1 was obtained by exchanging the eGFP gene in a peGFP-C1 construct. Therefore the mRFP1-gene was PCR-amplified from pRSET-B-mRFP1, a kind gift from Dr. Roger Y. Tsien (Howard Hughes Medical Institute, University of California at San Diego, La Jolla, CA), with two primers: 5'-GAATTCAGCGCTATGGCCTCCTCCGAGGACGTC (sense) and 5'- GAATTCAG ATCTGGCGCCGGTGGAGTGGCGG (antisense). The amplicon was subcloned in peGFP-C1 through the *Eco*47III and *Bgl*II restriction sites. The cloning of pmRFP-INs was similar; the mRFP1-gene was amplified with the same sense primer and with the 5'-GAATTCAGATCTGGGCGCCGGTGGAGTGGCGG antisense primer. The amplicon was subcloned in peGFP-IN^{s}-C2 (Maertens et al., 2003) to yield pmRFP-IN^{s}-C2. The clones were verified with PCR, *Pst*I-restriction and sequence analysis. pH1-eGFP was constructed by PCR-amplifying histone-H1 from an in-house construct, restriction with *Sal*I-*Bam*HI and ligation into peGFP-N1 restricted with *Xho*I-*Bam*HI. Clones were checked with PCR, restriction and sequencing analysis. The plasmid encoding eGFP-HP1β was a kind gift from Dr. Tom Misteli (National Cancer Institute, National Institutes of Health, Bethesda, MD,

USA). The cloning of the PWWP mutants of LEDGF/p75 was performed by Melissa McNeely in the Laboratory for Molecular Virology and Gene Therapy, under the supervision of Prof. Zeger Debyser.

4.2.2. Cell lines

HeLa cells were obtained from the NIH Reagent program and were grown in 'Complete Medium', high-glucose Dulbecco's Modified Eagle Medium (Gibco BRL, Belgium), supplemented with 10% heat-inactivated fetal bovine serum (Sigma-Aldrich gmbh, Taufkirchen, Germany) and 50 µg/mL Gentamycin (Gibco BRL, Belgium), at 5% CO_2 and 37 °C in a humidified atmosphere. To generate cell lines stably suppressing LEDGF/p75 mRNA, 20,000 HeLaP4-CCR5 cells were seeded in a 96-well plate and transduced with MLV-based retroviral vector encoding two identical miRNA-based hairpin sequences directed against the L3 sequence in the LEDGF/p75 mRNA (Maertens et al., 2003; Vandekerckhove et al., 2006) together with a Zeocin resistance cassette driven by a ubiquitous human immediate early CMV promoter. Cells were subsequently selected using Zeocin (Gibco BRL, Belgium) at 200 µg/ml. Selection resulted in a polyclonal cell line. A monoclonal cell line was generated by seeding at 0.2 cells per well and selection of a cell clone with strongest knockdown (HeLa-p75KD). The cell line stably suppressing LEDGF/p75 was made by Dr. Rik Gijsbers (Laboratory for Molecular Virology and Gene Therapy, K.U.Leuven). CHO cells stably expressing eGFP tagged hRPB1, the largest subunit of human RNA polymerase II, were a kind gift from Dr. Peter Cook (Sir William Dunn School of Pathology, University of Oxford, UK)(Sugaya et al., 2000).

4.2.3. Transfections

For the transient transfections 0.5-1×10^5 cells were seeded per well in a Lab-Tek™ Chambered Cover glass (VWR International, Leuven, Belgium) to obtain 50-70% cell confluency after overnight incubation. Transfections were performed with a Mirus TransIT®-HeLaMONSTER® transfection kit (VWR International, Leuven, Belgium), with 1 µL TransIT reagent, a maximum of 0.5 µg plasmid DNA per well and 1 µL MONSTER reagent per µg of DNA. After the transfection mixture was prepared in 50 µL fresh OptiMEM per well (Gibco BRL, Belgium), 450 µL prewarmed Complete Medium was added and this mixture was slowly added to the cells and the cells were incubated for at least 12 hours.

4.2.4. Western blotting, cellular fractionation assays

Western blotting was performed on whole cell lysates as described before (Vandekerckhove et al., 2006) with a specific anti-LEDGF/p75 antibody (A300-847A, Bethyl Laboratories, Montgomery, TX, USA). Cellular fractionation assays were performed as described before (Llano et al., 2006b), with the specific LEDGF/p75 antibody, an anti-RFP antibody for mRFP-IN (AB3216, Millipore N.V., Brussels, Belgium), an in-house polyclonal anti-GFP antibody and anti-α-tubulin (T5168, Sigma-Aldrich, Bornem, Belgium). Briefly, 24h post-transfection cells were harvested and either lysed for checking overall expression (T) or extracted with Triton X-100 to separate the soluble (S1) and insoluble (P1) cellular proteins. DNaseI-treatment allowed subsequent separation into a soluble, chromatin binding fraction (S2) and an insoluble, non-chromatin binding fraction (P2).

4.2.5. Cellular fluorescence correlation spectroscopy

For normal FCS measurements a commercial FCS/FCCS microscope (LSM510/ConfoCor2, Carl Zeiss, Jena, Germany) was used. Practically, enhanced green fluorescent protein (eGFP) (Heim et al., 1995) and monomeric red fluorescent protein (mRFP1) (Campbell et al., 2002) were used as fluorescent tags. The 488-nm line of the Ar^+-laser (acousto-optical tunable filter (AOTF) 0.1%, ~0.25 μW) was used to excite eGFP and the 543-nm line of the HeNe laser (AOTF 7%, ~8.6 μW) was used to excite mRFP1. The excitation light was reflected by a dichroic mirror (HFT 488/543) and focused through a C-Apochromat 40×/1.2 W Korr/0.13-0.17 objective. The fluorescence emission light was split by a second dichroic mirror (NFT 570) into two separate beam paths and passed through a 505-530 nm bandpass filter and 70-μm pinhole for eGFP fluorescence and a 600-650 nm bandpass filter and 78-μm pinhole for mRFP1 fluorescence. In each cell cytoplasm or nucleus, 10 consecutive fluorescence intensity (detected photons per second) and correlation measurements of 20 seconds were performed. The first fluorescence intensity curve was used to determine the photobleaching and the following nine correlation curves were averaged after individual inspection. All FCS measurements were performed at room temperature.

4.2.6. Cellular tunable focus FCS

For the cellular tunable focus FCS we used a home built FCS setup (Wawrezinieck et al., 2005)(Figure 4.1). The 30-mW 488-nm line of a Ar-Kr laser (Newport Spectra-Physics, Utrecht, The Netherlands) was split 50/50. One beam was maximally focused at the back-aperture of the objective (Olympus UPlanSApo 100x/NA1.4/Oil, Olympus Belgium N.V.,

Aartselaar, Belgium) to create wide-field excitation for observing the fluorescence of the cells through the eyepiece of the microscope (Olympus IX). The other beam was carefully expanded, collimated and directed centrically through an adjustable diaphragm, to allow for tunable focus FCS measurements. A fixed 50-µm pinhole was used for confocal detection on an avalanche photodiode (SPCM-AQR-15, Perkin-Elmer, Wiesbaden, Germany). The excitation power at each diaphragm setting was attenuated to give a constant power/area in the focal spot suitable for cellular FCS, avoiding potential intensity-dependent artefacts. Reproducible switching was achieved by measuring the residual power behind the diaphragm. These experiments were carried out at the Laboratory for Photochemistry and Spectroscopy, under the supervision of Prof. Johan Hofkens.

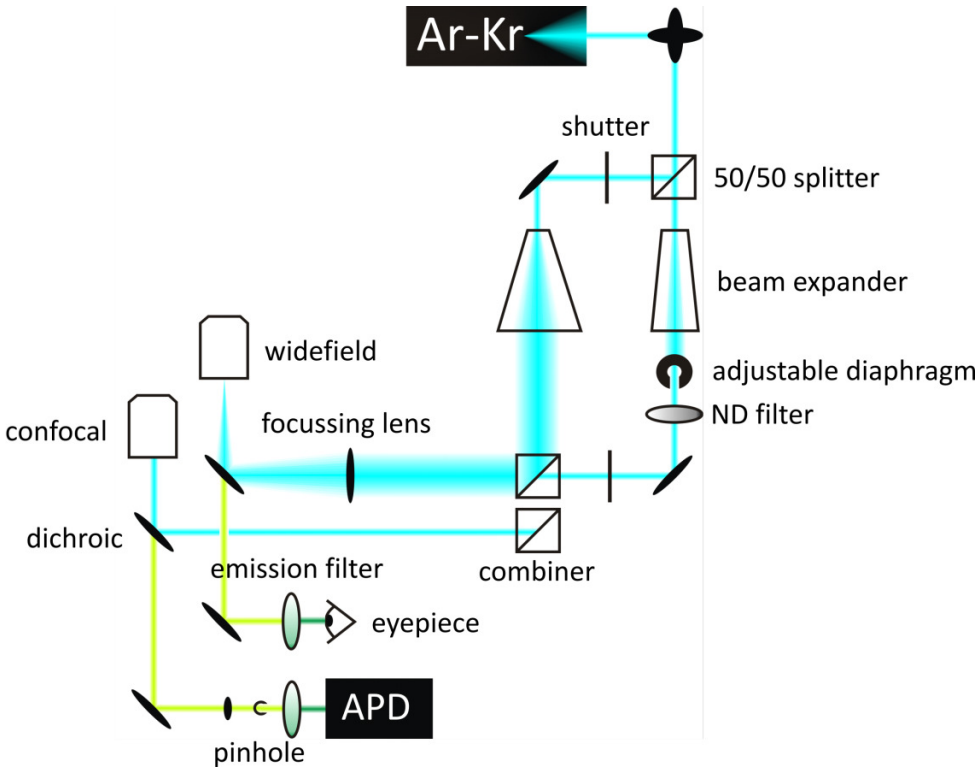

Figure 4.1 **Schematic representation of the tunable focus FCS setup** – Shutters controlled which path the excitation light followed. From the combiner either defocused (wide-field) or focused (confocal) excitation light travels to the objective. The focussing lens was removed for confocal excitation. Dichroic: dichroic mirror reflecting light below 490 nm. ND filter: attenuating neutral density filter. Emission filter : band pass 530/50. Pinhole: fixed at 50 µm. APD: avalanche photodiode.

4.2.7. FCS in the presence of binding to an immobile structure

Consider a binding equilibrium of a protein L that can interact with an immobile binding site S on the chromatin:

$$L + S \overset{K_{ass}}{\longleftrightarrow} LS \;\; with\, K_{ass} = \frac{k_{on}}{k_{off}}$$

$$association\, rate = k_{on}[L][S] \overset{S_0 \gg L_0}{\longrightarrow} (k_{on}S_0)[L] = k^*_{on}[L] \qquad \text{Equation 4.1}$$

$$dissociation\, rate = k_{off}[LS]$$

$$t_{free} = 1/k^*_{on} \;\; and \;\; t_{bound} = 1/k_{off}$$

With [L] the free concentration of protein L, in this case, LEDGF/p75, [S] the free concentration of immobile binding sites, [LS] the concentration of protein L bound to a binding site, K_{ass} the association binding constant, k_{on} the rate constant for association and k_{off} the rate constant for dissociation. When the concentration of binding sites is high, $k^*_{on} = k_{on}S_0$ is the pseudo rate constant for association. If L has a strong affinity for S, L will immobilize for a considerable amount of time after binding, which in term will lead to photobleaching of L during an FCS experiment and hence, the binding event will not be observed in the autocorrelation function. In terms of FCS, four dynamic regimes can generally be thought of:

4.2.7.1. Fast association ($t_{free} << \tau_{diff,free}$) and slow dissociation ($t_{bound} > \tau_{diff,free}$)

The protein binds to the immobile structure with very strong affinity. It will be photobleached in the laser focus when FCS is measured. FRAP is a better method to analyse these binding kinetics. FRAP will give a diffusion dependent slow to very slow recovery (Sprague and McNally, 2005).

4.2.7.2. Slow association ($t_{free} > \tau_{diff,free}$) and very slow dissociation ($t_{bound} > 100$ ms)

The chance of binding is low but once bound the fluorochrome will be photobleached. The fraction of the signal that photobleaches is a measure for the affinity:

$$\frac{F_{bleached}}{F_{unbleached}} = \frac{[LS]}{[L]} = \frac{k^*_{on}}{k_{off}} = K^*_{ass} \qquad \text{Equation 4.2}$$

K^*_{ass} (K_{diss}) can be determined by non-linear least-squares fitting of the percentage photobleaching versus the total concentration of protein L_0:

$$F_{bleached} = \frac{[LS]}{L_0} = y_0 + \frac{L_0 + S_0 + K_{diss} - \sqrt{(L_0 + S_0 + K_{diss})^2 - 4L_0S_0}}{2L_0} \quad \text{Equation 4.3}$$

In terms of LEDGF/p75 chromatin binding, $[LS]/L_0$ represents the fraction of eGFP-LEDGF/p75 that photobleaches. The total concentration L_0 is measured with FCS. The parameter y_0 is a constant accounting for the constant fraction of photobleaching at high concentrations.

4.2.7.3. Fast association ($t_{free} < \tau_{diff,free}$) and fast dissociation ($t_{bound} < \tau_{diff,free}$)

The binding reaction is observed as an additional slow component in the ACF, that is dependent on the size of the laser spot. The 'effective' diffusion coefficient defines the affinity:

$$D_{observed} = \frac{D_{free}}{1 + \frac{k^*_{on}}{k_{off}}} \quad \text{Equation 4.4}$$

4.2.7.4. Slow association ($t_{free} >> \tau_{diff}$), moderate dissociation ($\tau_{diff,free} < t_{bound} < 100ms$)

The binding reaction is observed as an additional slow component in the ACF, that is not dependent on the size of the laser focus. The fraction of free and 'apparent slow' diffusion components from the ACF is a measure for the affinity:

$$\frac{F_{slow}}{F_{fast}} = \frac{[LS]}{[L]} = \frac{k^*_{on}}{k_{off}} = K^*_{ass} \quad \text{Equation 4.5}$$

While in the first two regimes, the binding event is not seen in the ACF, in the third regime a seemingly slow diffusion is observed. In the fourth regime, as Michelman-Ribeiro et al. have shown, an FCS fit model for binding can be used (Michelman-Ribeiro et al., 2009):

$$G(\tau) = \frac{1}{2^{2/3}N}\left[F_{eq}G_D(\tau) + C_{eq}\exp^{-k_{off}t}\right]$$
$$F_{eq} = \frac{k_{off}}{k_{off} + k^*_{on}} \; and \; C_{eq} = \frac{k^*_{on}}{k_{off} + k^*_{on}} \quad \text{Equation 4.6}$$

with $G_D(\tau)$ the standard fit model (Equation 2.5) and F_{eq} the fraction of free protein and C_{eq} the fraction of bound protein. From this fit, both the diffusion coefficient of the free species as the rate constants for association and dissociation can be determined.

4.2.8. Scanning confocal fluorescence recovery after photobleaching

Fluorescence recovery after photobleaching (FRAP) is a quantitative fluorescence technique that is used to study protein dynamics (Axelrod et al., 1976). Scanning confocal FRAP is performed on a laser scanning microscope. By illuminating a defined region-of-interest in the sample with a brief high-intensity laser pulse, the fluorescence is rapidly photobleached. It will recover due to the exchange of bleached molecules inside with unbleached molecules outside this region. The fluorescence recovery is generally determined by the diffusion rate of the molecules under investigation, but can also be determined by binding reactions to slow or immobile structures. In this research, FRAP was used to investigate whether or not immobile cellular subpopulations existed and to investigate dynamics that were too slow to be studied with FCS. Cellular FRAP curves for eGFP-LEDGF/p75 were obtained on a LSM510 (Carl Zeiss). A circular bleach spot in the nucleus with a radius of 1 μm was briefly scanned once with a high laser intensity (AOTF 100%) and the average fluorescence intensity (in arbitrary units) was subsequently monitored at low laser intensity (AOTF 2%) as a function of time. Neglegible acquisitional photobleaching (as measured for eGFP) was observed and measurements were normalized before averaging different cells. The optical path was similar to that for FCS. Depending on the nature of the dynamic process, different fit models have to be used to analyze FRAP data.

4.2.8.1. FRAP in the case of Brownian diffusion

A relative simple closed form solution of the diffusion equation exists for FRAP when measurements are performed in a circular disk profile (Axelrod et al., 1976; Soumpasis, 1983; Sprague et al., 2004):

$$fluo(t) = \exp\left(-\frac{\tau_{diff}}{2t}\right)\left[I_0\left(\frac{\tau_{diff}}{2t}\right) + I_1\left(\frac{\tau_{diff}}{2t}\right)\right](1 - F_{imm}) \quad \text{Equation 4.7}$$

With τ_{diff} the characteristic diffusion time of the probe through the bleach spot, F_{imm} the fraction of the fluorescence intensity that does not recover and I_α a modified Bessel function of the first kind and order α. This analysis model can be approximated by a Taylor expansion for fast recoveries (Figure 4.2):

$$I_0\left(\frac{\tau_{diff}}{2t}\right) + I_1\left(\frac{\tau_{diff}}{2t}\right) = 1 + \frac{1}{2}\left(\frac{\tau_{diff}}{2t}\right) + \frac{1}{4}\left(\frac{\tau_{diff}}{2t}\right)^2 + \frac{1}{16}\left(\frac{\tau_{diff}}{2t}\right)^3 + \frac{1}{64}\left(\frac{\tau_{diff}}{2t}\right)^4 + \frac{1}{348}\left(\frac{\tau_{diff}}{2t}\right)^5 \quad \text{Equation 4.8}$$

The diffusion coefficient from a FRAP measurement can be calculated as follows:

$$\tau_{diff} = \frac{\omega^2}{D} \qquad \text{Equation 4.9}$$

where ω^2 is the radius of the bleach spot.

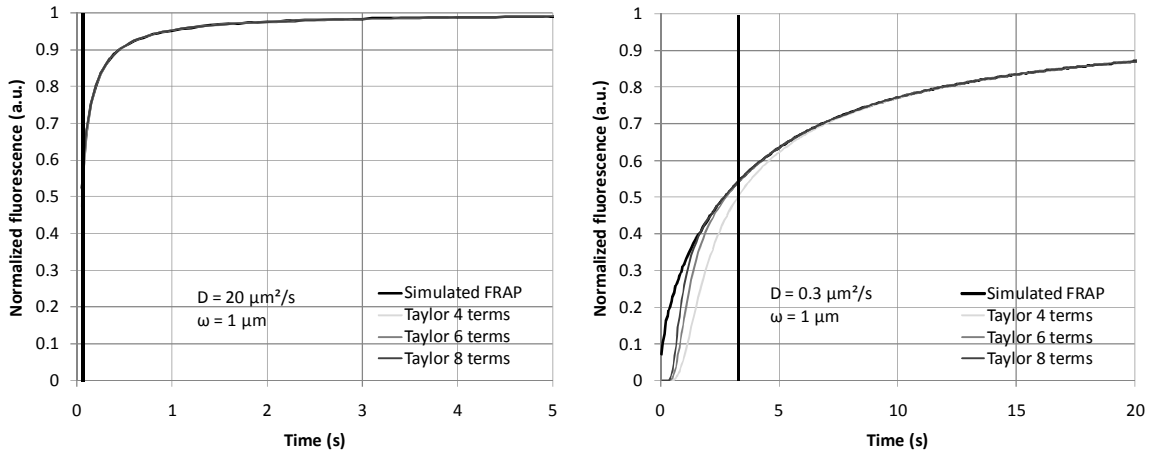

Figure 4.2 **Simulated FRAP curve for a fast recovery (left) and a slow recovery (right)** The respective approximations by a 4, 6 or 8 term Taylor expansion of the sum of Bessel functions. In the case of fast diffusion, practically the whole FRAP curve is described well with any Taylor expansion. However, in the case of slow diffusion, a significant part of the FRAP curve (left of the black vertical line) is not described well with the Taylor expansion.

As a rule of thumb, the circular disk method generates accurate diffusion coefficients as long as the bleach pulse is not longer than one tenth of the diffusion time (Meyvis et al., 1999). More complex fit models for studying diffusion with FRAP exist that take the 3-dimensional shape of the bleach spot into account, but since FCS can be used for studying diffusion accurately we only mention this as a reference (Braeckmans et al., 2003). Equation 4.7 can be modified to account for anomalous diffusion (Feder et al., 1996).

4.2.8.2. FRAP in the case of binding to an immobile structure

4.2.8.2.1. Binding kinetics – diffusion uncoupled FRAP

When a protein can bind strongly to an immobile structure, the FRAP recovery will be governed by the kinetics of this binding reaction and can be accurately described by an exponential function, or in the case of multiple binding sites, a sum of exponentials (Sprague et al., 2004):

$$frap(t) = 1 - \sum_{i=1}^{n} \frac{k^*_{on,i}}{k_{off,i} + k^*_{on,i}} \exp(-k_{off,i}t) \qquad \text{Equation 4.10}$$

Where k^*_{on} is the pseudo rate constant for association, k_{off} the rate constant for dissociation and n the total number of binding sites. In this case, k^*_{on} follows from the maximum possible bleach depth. If $k^*_{on} >> k_{off}$, then the maximum bleach depth is zero.

4.2.8.2.2. Binding kinetics – diffusion coupled FRAP

When a diffusion model (Equation 4.7) gives a good fit, but a much lower D is obtained than expected (relative to a reference protein), binding kinetics are likely contributing to the recovery. When the relaxation time for the association process, $1/k^*_{on}$, is smaller than the average diffusion time, diffusion will still contribute to the FRAP recovery, and Equation 4.10 will not describe the dynamics correctly. In principle, even very slow (small k_{off}) FRAP recoveries can still be governed by diffusion (large k^*_{on}). The observed diffusion coefficient can be used to calculate the ratio of rate constants, the pseudo equilibrium constant. The existence of different binding sites cannot be inferred from such a recovery, since only the sum of the ratios is obtained (Sprague et al., 2004):

$$D_{observed} = \frac{D_{free}}{1 + \sum_{i=1}^{n} \frac{k^*_{on,i}}{k_{off,i}}} \qquad \text{Equation 4.11}$$

When $1/k^*_{on}$ and τ_{diff} are in the same order of magnitude, a more complex FRAP model has to be used, that includes both rate constants and the diffusion coefficient as fit parameters (Sprague et al., 2004).

4.2.8.2.3. Discerning diffusion coupled from uncoupled FRAP

In general, the contribution of diffusion to the observed FRAP recovery can be checked by varying an experimental parameter that only influences the diffusion time scale. For example:

- If the FRAP measurement is performed in a spot with variable size, the recovery should be slower for a bigger spot if diffusion contributes to the observed dynamics. If only k_{off} determines the FRAP, the curves should be identical.
- Similarly, if a half-nucleus bleaching is performed and the recovery is measured at different distances from the bleached region, the curves should be identical if only k_{off} determines the recovery.
- Finally, when a half-nucleus bleaching is performed and the recovery is measured on a line orthogonal to the bleach border, the line profile has a different evolution if diffusion or binding determines the dynamics (Figure 4.3) (Mueller et al., 2008).

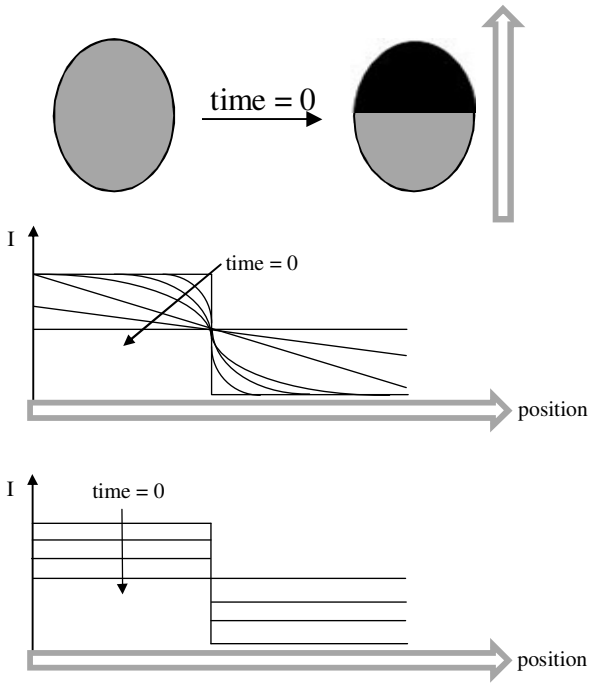

Figure 4.3 **Half nucleus bleaching to check diffusion dependence** (top) Half of a nucleus is photobleached instantaneously and the recovery is measured on a line orthogonal to the bleach border. (middle-bottom) Evolution of the fluorescence along the arrow when diffusion describes the dynamics (middle) or when the binding kinetics are rate limiting (bottom).

4.2.9. Structure based mutation of PWWP residues important for DNA binding

To determine the residues important for the interaction of the PWWP domain of LEDGF/p75 with DNA we used biomolecular modelling. First, we modelled the structure of the PWWP domain of LEDGF/p75 using the latest available version of Modeller (Sali and Blundell, 1993), with the NMR structure (PDB 2B8A) of the HDGF-PWWP as a template (Lukasik et al., 2006). Secondly, we predicted putative DNA binding residues using the HotPatch algorithm (Pettit et al., 2007). Among these we chose the residues K56 and R74 that were solvent exposed and not required to stabilize the tertiary structure of the protein. In this way, we avoided the generation of a mutant protein that would be unable to form the correct PWWP-fold. Of concern, these residues in the LEDGF/p75-PWWP were identical to the residues in the HDGF-PWWP sequence and had already been shown to be important for DNA binding of the HDGF-PWWP (Lukasik et al., 2006). The modelling part of this chapter was performed by Arnout Voet at the Laboratory for Biomolecular Modelling at K.U.Leuven.

4.3. Results

4.3.1. Subpopulations of LEDGF/p75 in the living cell

To image LEDGF/p75 in the living cell, we expressed a fusion protein of LEDGF/p75 and the enhanced green fluorescent protein (eGFP), eGFP-LEDGF/p75, displaying a heterogeneous nuclear distribution characteristic of endogenous LEDGF/p75 (Figure 4.4 A) (Cherepanov et al., 2003). Western blotting revealed no protein degradation of eGFP-LEDGF/p75 (Figure 4.4 B). Chromatin binding properties of the fusion were similar to that of endogenous LEDGF/p75 as demonstrated by *in vitro* cell fractionation (Figure 4.4 C): following extraction of the cells with Triton X-100, eGFP-LEDGF/p75 was present in the insoluble pellet (P1), whereas a DNaseI treatment of the P1 pellet released the protein into the supernatant (S2), as demonstrated earlier for endogenous LEDGF/p75 (Llano et al., 2006b). Thus, a N-terminal fluorescent tag does not affect the chromatin binding properties of LEDGF/p75. All further experiments were carried out in LEDGF/p75 knockdown HeLa cells (Figure 4.4 D) with transiently expressed eGFP-LEDGF/p75. In a spot in the nucleoplasm with bright and homogenous eGFP fluorescence we measured the fluorescence intensity I_{fluo} in function of time. While I_{fluo} instantaneously decreased (~40%) for eGFP-LEDGF/p75 (Figure 4.4 E, red curve), it remained largely constant for our control protein eGFP (Figure 4.4 E, black curve). At steady state, I_{fluo} in the focal spot of the laser is a measure for the dynamic equilibrium between molecules in the spot that have a probability of being photobleached by the laser and unbleached molecules outside the spot. We hypothesized that eGFP-LEDGF/p75 binds to chromatin, which increases the residence time in the focal spot and hence the photobleaching probability. Photobleaching was concentration-dependent in the nM range (Figure 4.4 F), we determined a concentration of 'chromatin binding sites' of 50 nM and a steady state dissociation constant of 100 nM (Equation 4.3). The residual constant fluorescence intensity after the initial photobleaching was used to calculate an autocorrelation function (ACF), using Fluorescence Correlation Spectroscopy (FCS). This allowed us to study the dynamics of the bleaching insensitive fraction of eGFP-LEDGF/p75. Our control, eGFP, exhibited free diffusion in the nucleus (Figure 4.4 G, red ACF) and a diffusion coefficient (D) of 21.8 $\mu m^2/s$ was calculated (Table 4.1), which is in good agreement with previous findings (Chen et al., 2002).

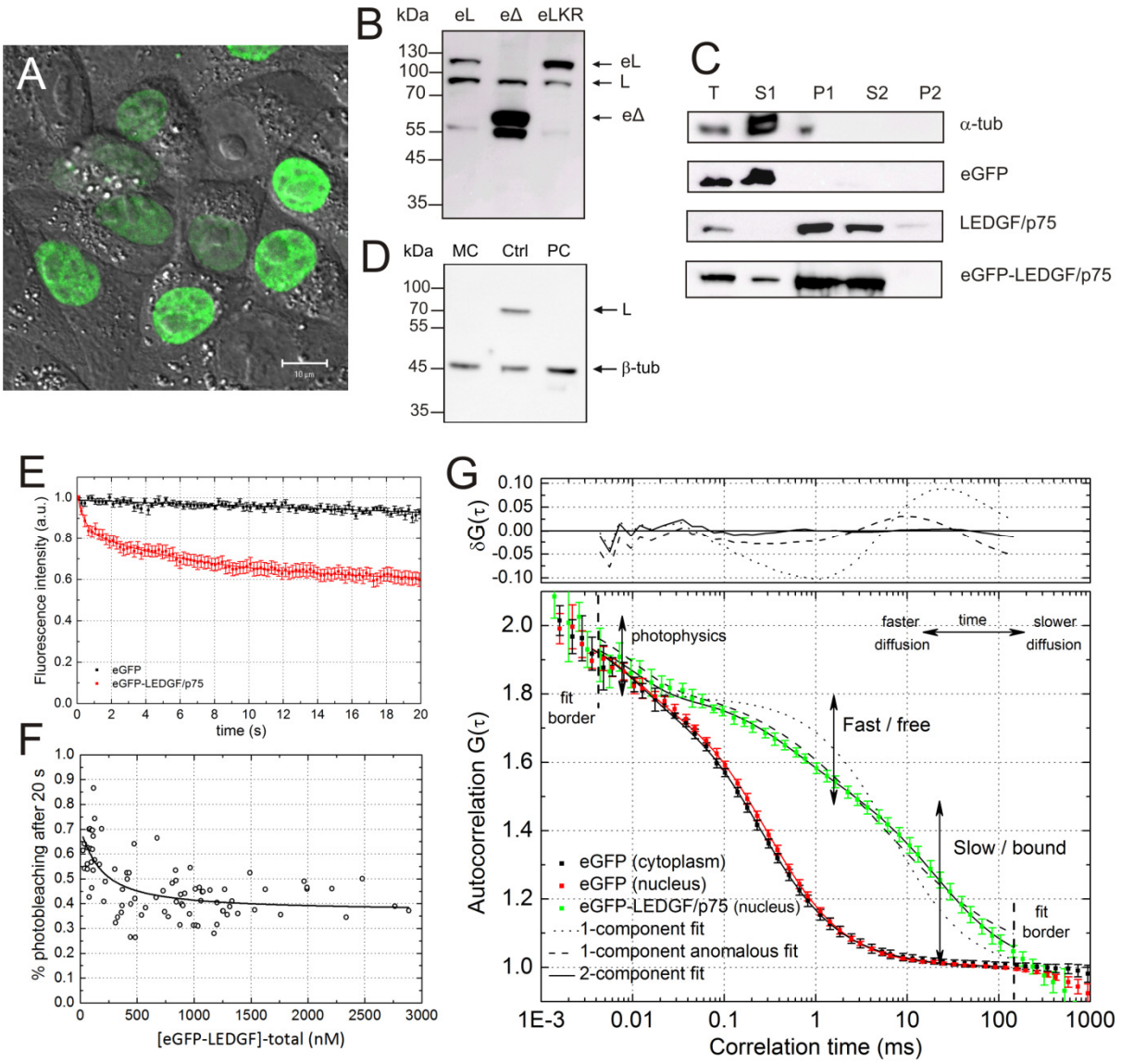

Figure 4.4 **Dynamic subpopulations of LEDGF/p75 revealed by photobleaching and FCS** (A) Confocal fluorescence image of HeLa cells expressing eGFP-LEDGF/p75 with a transmission image overlay, obtained by DIC microscopy. Scale bar = 10 μm. (B) Western blot with an anti-p75 antibody of HeLa cells transiently expressing eGFP-fusions. eL = eGFP-LEDGF/p75, eΔ = eGFP-Δ325, KR = K56D-R74D, L = endogenous LEDGF/p75. (C) Cellular fractionation assay of HeLa cells transiently expressing eGFP-fusions. Western blot of different fractions is shown using antibodies to indicated proteins. T = total cell lysate, S1 = Triton-soluble cellular fraction, P1 = Triton-insoluble cellular fraction, S2 = DNase/$(NH4)_2SO_4$-soluble cellular fraction, P2 = DNase/$(NH4)_2SO_4$-insoluble cellular fraction. (D) Western blot showing strong knockdown of endogenous LEDGF/p75 in HeLa-p75KD cells. MC = monoclonal HeLa-p75KD cells, Ctrl = control HeLa cell line, PC = polyclonal knockdown HeLa cell line. Equal loading was verified with an antibody against β-tubulin. L = endogenous LEDGF/p75. (E) Fluorescence intensity of eGFP and eGFP-LEDGF/p75 in living HeLa cells. (F) Concentration dependence of photobleaching of eGFP-LEDGF/p75 reveals stronger affinity sites. Solid line = fit with Equation 2.20. (G) Normalized ACFs of eGFP and eGFP-LEDGF/p75. The upper panel is a residual plot showing the best fit with a two-component diffusion model. Error bars = s.d.

The eGFP-LEDGF/p75 ACF on the other hand was strongly shifted to a slower time scale (Figure 4.4 G, green ACF), and could only be described using a 2-component model (compare $\delta G(\tau)$ in Figure 4.4 H): a minor component (35%) having an apparent D of 14.6 $\mu m^2/s$, in accordance with the value for free diffusion (14.7 $\mu m^2/s$), and a major component (65%) having a very low mobility (D = 0.3 $\mu m^2/s$) (Table 4.1). We hypothesized that even this slow component of the bleaching insensitive fraction of the LEDGF/p75 population involves interactions with chromatin. In the following two sections we will explain these observations further.

Table 4.1 **Subcellular protein dynamics of LEDGF/p75 and variants measured with FCS** - Intracellular ACFs were fitted with a two-component or anomalous diffusion model (Equation 2.20) and the fraction (F) and diffusion coefficients (D, in $\mu m^2/s$) was calculated as described in the experimental procedures. The calculations of D were made assuming a diffusion coefficient of Rhodamine 6G equal to 280 $\mu m^2/s$. The K150A mutant of LEDGF/p75 is explained in detail in Chapter 5. mNLS = mutated nuclear localisation signal, mPWWP = mutated PWWP domain, n = number of independent cellular measurements, s.d. = standard deviation, n.p. = fit not possible, * = fit not satisfying, a.u. = arbitrary units.

Protein	properties	n	D_{fast} ± s.d.	F_{fast} ± s.d.(%)	D_{slow} ± s.d.	α ± s.d.(a.u.)
			2-component fit			anomalous fit
Nucleus						
eGFP-LEDGF/p75	wild type	23	14.6 ± 2.6	35.9 ± 3.0	0.3 ± 0.1	n.p.
eGFP-LEDGF/p75 K150A	mNLS	13	13.4 ± 2.2	37.6 ± 4.5	0.3 ± 0.1	n.p.
eGFP-LEDGF/p75 K56D	mPWWP	22	10.8 ± 2.3	44.5 ± 9.3	0.7 ± 0.2	0.69 ± 0.05*
eGFP-LEDGF/p75 R74D	mPWWP	21	8.8 ± 1.6	58.5 ± 8.6	0.9 ± 0.4	0.78 ± 0.04
eGFP-LEDGF/p75 K56D R74D	mPWWP	28	9.9 ± 1.9	66.8 ± 9.6	1.0 ± 0.5	0.80 ± 0.05
eGFP-Δ325		9	14.0 ± 1.3	89.7 ± 2.8	0.6 ± 0.4	0.88 ± 0.03
eGFP		19	21.8 ± 2.4	97.3 ± 0.7	0.5 ± 0.3	0.98 ± 0.02
Cytoplasm						
eGFP-LEDGF/p75 K150A	mNLS	18	11.9 ± 1.7	53.3 ± 9.4	1.0 ± 0.4	0.78 ± 0.03
eGFP-Δ325		10	16.9 ± 1.1	91.9 ± 4.3	1.0 ± 0.6	0.93 ± 0.05
eGFP		17	23.8 ± 1.2	93.7 ± 2.8	0.6 ± 0.4	0.93 ± 0.05

4.3.2. Dynamic chromatin interaction of LEDGF/p75

We constructed an eGFP fusion of the C-terminal Δ325-fragment of LEDGF/p75 that lacks the domains important for chromatin and DNA binding (Figure 1.8). The photobleachable fraction of this fragment in the nucleus decreased to that of freely diffusing eGFP (Figure 4.5 A, red curve) and the ACF showed that the large majority (~90%) exhibited free intranuclear diffusion (Figure 4.5 B, red curve and Table 1). Chromatin binding of LEDGF/p75 was thus indeed responsible for the photobleaching and the slow component of the ACF. Next, we compared our measurements with those of well characterized chromatin interacting proteins: Heterochromatin Protein 1 (HP1), RNA polymerase II (RNAPII) and Linker Histone H1 (H1). HP1 is a protein involved in heterochromatin formation through interaction with and oligomerization on nucleosomes (Eissenberg and Elgin, 2000). RNAPII is a multi-protein complex that synthesizes messenger RNA (Sugaya et al., 2000). H1 binds very strongly to the crossing-over of DNA twisted around an octamer of core-histone proteins stabilizing the nucleosome (Zlatanova et al., 2000).

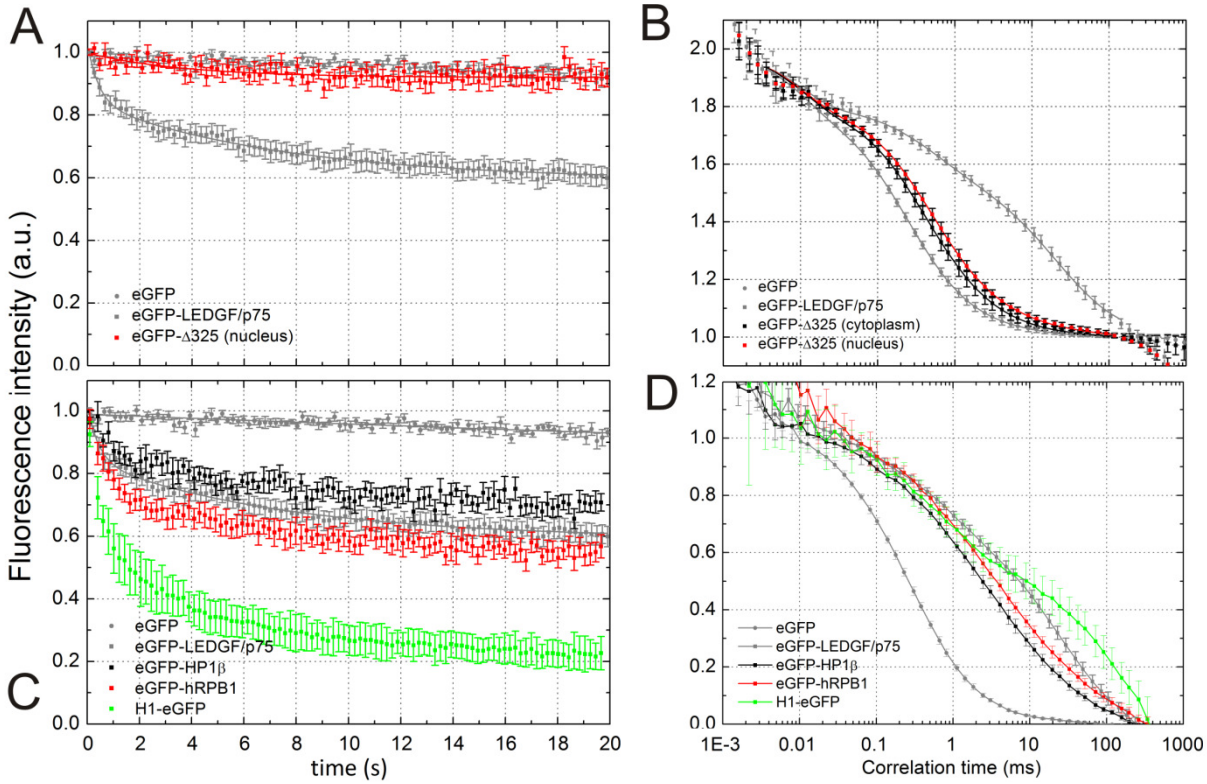

Figure 4.5 **Dynamic chromatin interaction of LEDGF/p75** (A) Cellular point fluorescence intensity and (B) ACF of eGFP-Δ325. Measurements on eGFP-LEDGF/p75 and eGFP are also plotted in gray as a reference. (C) Cellular point fluorescence intensity and (D) ACF of eGFP-HP1β, RNA polymerase II (CHO cells stably expressing eGFP-hRPB1) and Histone H1-eGFP. Error bars = s.d.

All three proteins were expressed as a fusion with eGFP. For HP1 and RNAPII we observed markedly more photobleaching than for eGFP (Figure 4.5 C, resp. black and red curves), although their ACFs demonstrated that both proteins have a slightly higher mobility than LEDGF/p75 (Figure 4.5 D). For H1 we observed very strong photobleaching that increased throughout the measurements, distorting the ACF considerably (Figure 4.5 C and D, green curves). In conclusion, chromatin interacting proteins generally diffuse slowly, as measured with FCS and are hence more sensitive to photobleaching.

4.3.3. Possible explanations for the observed dynamics

Five mechanisms of chromatin binding of LEDGF/p75 can be thought of: (1) LEDGF/p75 binds very strongly to chromatin and the observed diffusion coefficient (D_{obs} = 0.3 μm²/s, $\tau_{diff,obs}$ = 25.9 ms) would reflect diffusion of chromatin (Figure 4.6 A). This is unlikely, since reported values for the diffusion coefficient of chromatin are 2-3 orders of magnitude lower than what we observed (Marshall et al., 1997).

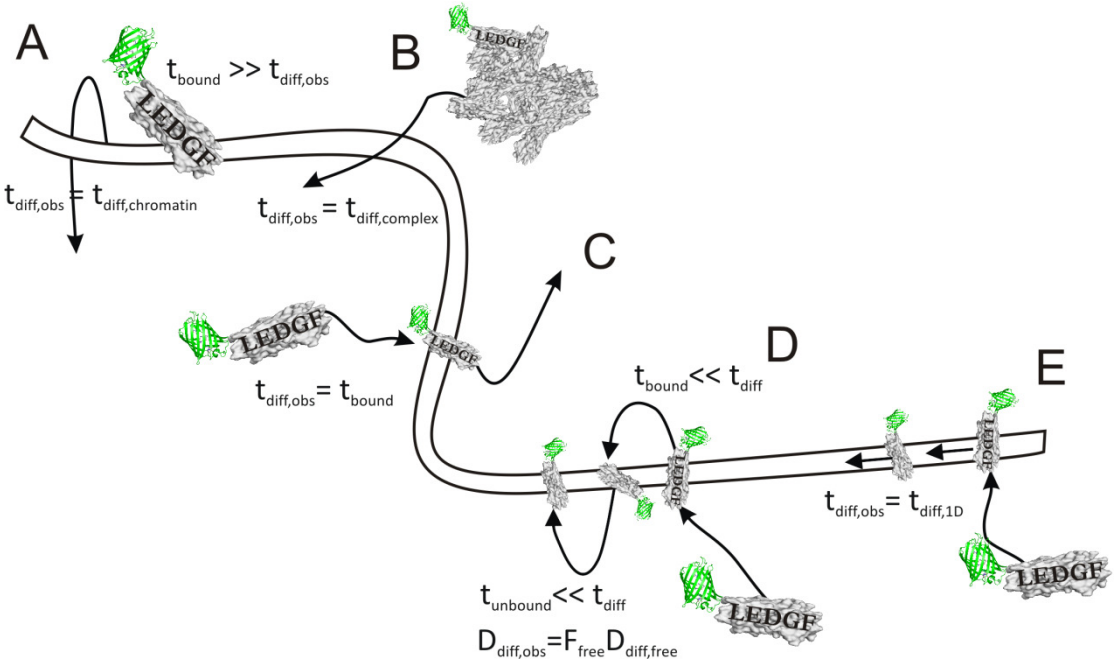

Figure 4.6 **Five possible explanations for the slow dynamics of LEDGF/p75** (A) The observed diffusion time t_{diff} is that of the chromatin if LEDGF/p75 strongly associates with the chromatin. t_{bound} is the bound time. (B) A multiprotein complex exhibits a very slow t_{diff}. (C) Diffusion-uncoupled FCS when $t_{bound} = t_{diff}$. $t_{unbound}$ is the unbound time and is dictated by the fraction of the fast (F_{free}) and slow (F_{bound}) components in the FCS curve. (D) Reaction-diffusion-coupled FCS when both $t_{unbound}$ and t_{bound} are smaller than t_{diff}. The ratio of $t_{diff,obs}$ and the t_{diff} for free diffusion dictates F_{free} and F_{bound}. (E) One-dimensional diffusion on the chromatin.

(2) LEDGF/p75 is part of a multi-protein complex and the observed diffusion coefficient is a measure for the size of the complex (Figure 4.6 B). This is also unlikely, since a macromolecule (>MDa) would suffer from obstructed (anomalous) diffusion (Weiss et al., 2004), which was not observed (Figure 4.4 H). Moreover chromatin binding is definitely contributing strongly to the observed dynamics (Figure 4.5 A-D). (3) LEDGF/p75 is immobilized during 25.9 ms ($t_{diff,obs}$= the bound time t_{bound}) after associating with chromatin. In this case diffusion would not contribute to the observed dynamics (Figure 4.6 C). (4/5) LEDGF/p75 binds faster to the chromatin than the diffusion time, in a transient (Figure 4.6 D) and/or continuous way (Figure 4.6 E). In this case the binding reaction but also diffusion would contribute to the observed dynamics.

4.3.4. Dynamic chromatin scanning by LEDGF/p75

We used an in-house developed tunable focus fluorescence correlation spectroscope (TF-FCS) adapted for cellular measurements to determine whether the ACF was affected by the size of the confocal laser spot and thus by diffusion (see section 4.2.6 for an optical scheme of the setup). *In vitro* control measurements of a standard probe rhodamine 6G in MQ (Figure 4.7 A-C) showed that it was indeed possible to tune the time scale of diffusion in a broad range. Furthermore, measurements on eGFP in PBS (Figure 4.7 D) and of eGFP in a buffer with a refractive index matching that of the intracellular environment (PBS, 23.5% w/w sucrose, n=1.37) showed that the setup also performed well for measurements on eGFP under optical conditions similar to the intracellular environment.

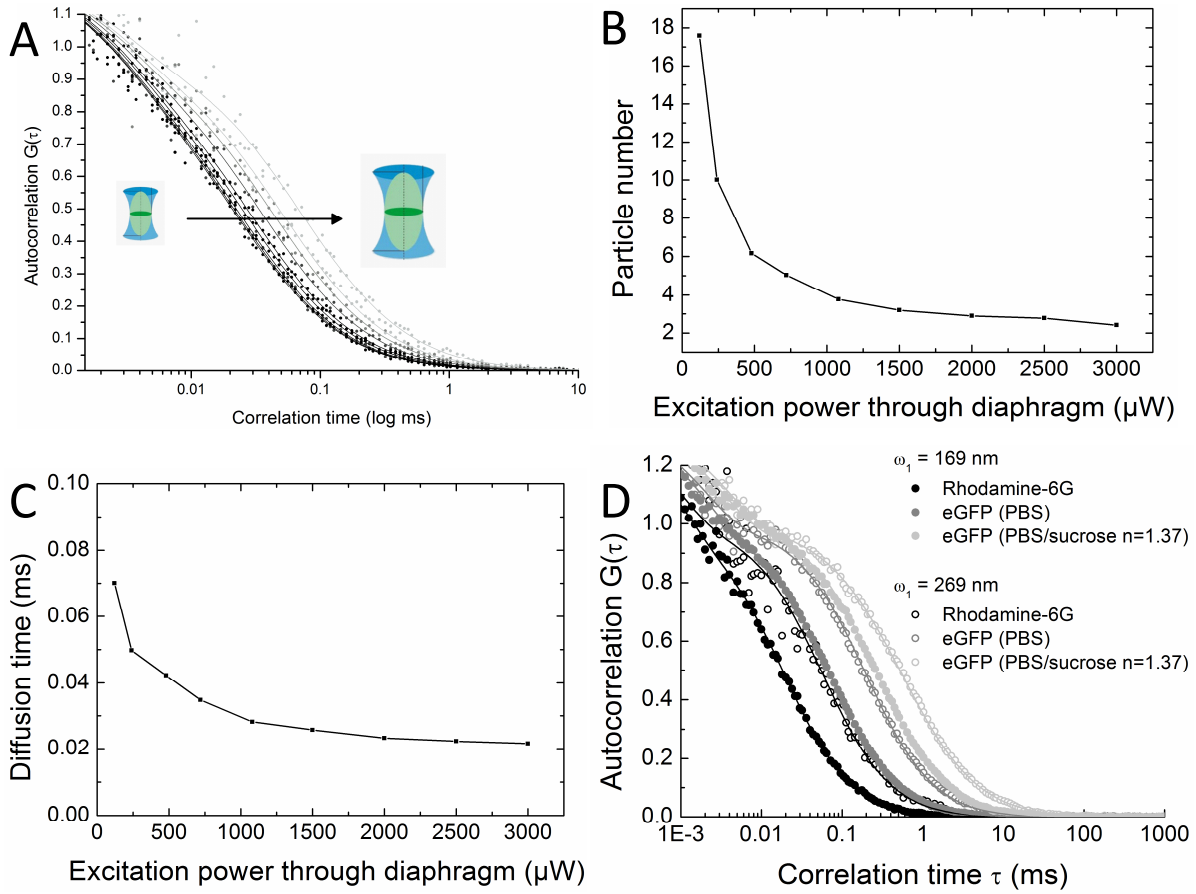

Figure 4.7 **Performance of the tunable focus FCS setup** (A) When the focus is enlarged, the experimental ACF of rhodamine 6G shifts to a slower time scale. When the diaphragm is opened, more collimated light goes through, as measured with a power meter. Both the particle number (B) and the diffusion time (C) of rhodamine 6G decrease when the beam is enlarged, i.e. when the focus is smaller. (D) The *in vitro* ACFs of rhodamine 6G in MQ, eGFP in PBS and eGFP in PBS/sucrose all shift towards a slower time scale when the focus is enlarged. ω_1 = radial radius of the focus.

In cells, the ACF of intracellular eGFP was also clearly dependent on the size of the focal spot (Figure 4.8, black curves), consistent with the purely Brownian diffusion of the protein. The ACF of eGFP-LEDGF/p75 in the cell nucleus was also shifted in a larger focal spot (Figure 4.8, red curves), especially for the slower component of the curve. The slow component in the ACF is thus dependent on diffusion, and argues for an extraordinary dynamic interaction of LEDGF/p75 with chromatin (Figure 4.6 D and/or E).

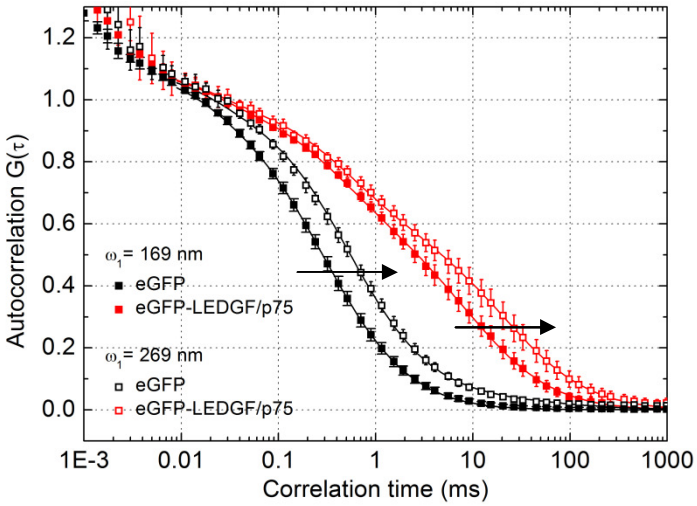

Figure 4.8 **Diffusion-controlled fast genome scanning by LEDGF/p75** – Cellular tunable focus fluorescence correlation spectroscopy measurements of eGFP and eGFP-LEDGF/p75 in HeLa cells. ω_1 is the radial radius of the excitation focus, experimentally determined from *in vitro* calibration measurements. Error bars = s.d.

If a binding reaction to an immobile reactant such as the chromatin rapidly reaches an equilibrium, diffusion will appear to be slowed down. The 'effective' diffusion coefficient will be $D_{obs,eff} = D_{free}F_{free}$ with D_{free} the expected diffusion coefficient for free nuclear diffusion (= 14.7 μm²/s for eGFP-LEDGF/p75) and F_{free} the average fraction of molecules not bound to chromatin (Crank, 1975). Based on our measured $D_{obs,eff}$ = 0.3 μm²/s LEDGF/p75 remains 98% of the time bound to chromatin and during the remaining 2% of the time, LEDGF/p75 quickly moves to a new binding site. Next, it follows from a simple protein binding equilibrium that $k^*_{on}/k_{off} = (1-F_{free})/(F_{free})$. The pseudo rate constant for association k^*_{on} (= k_{on} multiplied with the concentration of binding sites) for LEDGF/p75 is thus 49 times larger than the rate constant for dissociation k_{off}. Finally, since the binding times $t_{unbound} = 1/k^*_{on}$ and $t_{bound} = 1/k_{off}$ would have to be faster than the expected diffusion time (t_{diff} = 25.9 ms), the average k_{off} (the rate constant for dissociation) would have to be larger than 38.6 s^{-1} and k^*_{on} larger than $1.9x10^3$ s^{-1}. These estimations are in line with an effective diffusion regime, previously simulated for the analysis of FRAP and FCS measurements (Sprague et al., 2004; Michelman-Ribeiro et al., 2009). An association purely based on a collision of LEDGF/p75 and chromatin would result in a $k_{on} \sim 10^9$ $M^{-1}s^{-1}$. We estimated k^*_{on} larger than $1.9x10^3$ s^{-1}, so we estimate the concentration of binding sites of LEDGF/p75 on chromatin to be larger than 1.9 μM. If we assume a globular HeLa nucleus with a 10-μm diameter, 1.9 μM of binding sites roughly corresponds to 10^6 binding sites. This

concentration is consistent with our observation that the parameters from our FCS measurements were constant in the measured concentration range of LEDGF/p75 (< 3 μM). Importantly, at nM concentrations we observed stronger chromatin binding (Figure 4.4 E), suggesting the existence of a few (50 nM) strong-affinity binding sites. In summary, since the observed dynamics of chromatin binding scale with the confocal volume size, the association and dissociation time scales are smaller than the diffusion time scale, the apparent slow dynamics in our FCS measurements are thus due a high frequency of binding to chromatin, (high k^*_{on}, high k_{off}). Strong-affinity chromatin binding (mechanism 3) thus does not determine the dynamics of LEDGF/p75 in the living cell nucleus. Rather, LEDGF/p75 predominantly moves by dynamic 'genome scanning' (mechanism 4/5), a fast weak-affinity diffusion-collision driven chromatin interaction.

4.3.5. Chromatin scanning of LEDGF/p75 is decelerated by HIV-1 integrase

We checked the chromatin interaction of eGFP-LEDGF/p75 in absence or presence of IN. In absence of mRFP-tagged IN, about 40% of the fluorescence signal of eGFP-LEDGF/p75 photobleached during the first seconds. Upon co-expression of mRFP-tagged IN, this photobleaching increased considerably (Figure 4.9 left). A larger fraction of the LEDGF/p75 population photobleached (on average 75% instead of 40%), similar to the photobleaching observed for histone H1, a strong chromatin binding protein. These experiments imply an increased affinity of LEDGF/p75 for chromatin, upon co-expression of IN. Moreover, at increasing concentrations of eGFP-LEDGF/p75 a stronger interaction with chromatin was observed (Figure 4.9 right), resulting in up to 90% photobleaching during the first 20 seconds. Importantly, in cells expressing only IN, this photobleaching was not present (Figure 5.3 and (Maertens et al., 2005)).

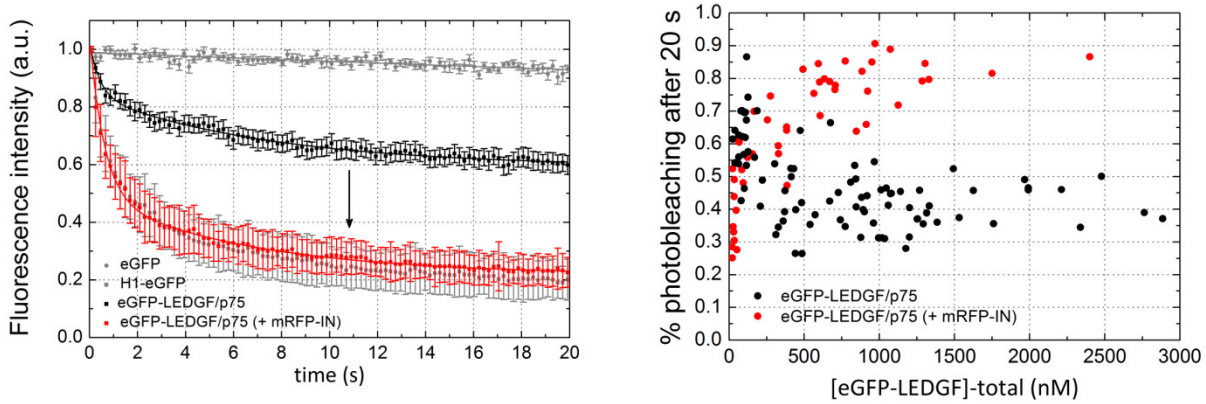

Figure 4.9 **Dynamic chromatin scanning of LEDGF/p75 is decelerated by HIV-1 integrase.** (left) Point fluorescence measurement of eGFP-LEDGF/p75 in cells co-expressing mRFP-IN. (right) Concentration-dependence of photobleaching of eGFP-LEDGF/p75.

In order to estimate the residence time of eGFP-LEDGF/p75 on chromatin in the presence of mRFP-IN, we used fluorescence recovery after photobleaching (FRAP). As expected, no permanent immobile fraction was observed in cells only expressing eGFP-LEDGF/p75, reaching 90% recovery in ~10 seconds (Figure 4.10 left, black curve), which is consistent with a protein that readily dissociates from the chromatin. Upon co-expression of mRFP-IN however, the fluorescence recovery was an order-of-magnitude slower and incomplete (Figure 4.10 left, red curve), reaching only about 60% after 100 seconds.

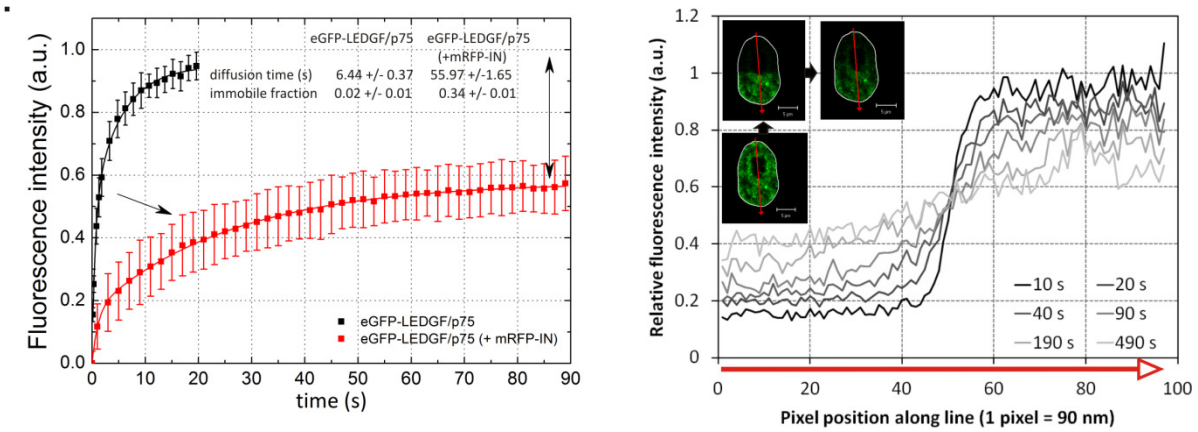

Figure 4.10 **Dynamic chromatin scanning of LEDGF/p75 is decelerated by HIV-1 integrase.** (left) Cellular FRAP measurement of eGFP-LEDGF/p75 in cells expressing mRFP-IN. Error bars = s.d. (right) Control photobleaching experiment to verify the contribution of diffusion to the observed photobleaching recovery (Mueller et al., 2008).

We also used a different FRAP method, to check the contribution of diffusion to the observed dynamics, analogously as we did for the fast dynamics of eGFP-LEDGF/p75 in absence of mRFP-IN. When photobleaching half of the nucleus, the fluorescence at a certain position either recovers with the same rate irrespective of the distance from the unbleached area (diffusion uncoupled mechanism) or exhibits a distance dependent recovery (diffusion coupled mechanism)(Mueller et al., 2008). This method is explained in more detail in the Materials and Methods section. Clearly, LEDGF/p75-IN still exhibited diffusion coupled recovery (Figure 4.10 right). Not surprisingly thus, a standard diffusion model (Axelrod et al., 1976) described the FRAP data in Figure 4.10 nearly perfectly, even though the recovery of LEDGF/p75-IN is shifted by almost an order-of-magnitude to a slower time scale with respect to LEDGF/p75 alone.

In conclusion, upon co-expression of IN, the dynamics of LEDGF/p75 shift towards a tight chromatin association. In absence of LEDGF/p75, IN behaves as a free protein (section 5.3.1) and in absence of IN, LEDGF/p75 shows moderate affinity for chromatin (Figure 4.4 H). Upon their co-expression, the protein population segregates into a concentration-dependent immobile fraction and a mobile fraction with order-of-magnitude slower dynamics.

4.3.6. The PWWP domain contributes to strong affinity chromatin binding of LEDGF/p75

Finally, we set out to verify the previously documented contribution of the PWWP-domain to the overall chromatin binding of LEDGF/p75 (Turlure et al., 2006; Llano et al., 2006b). To affect the overall protein structure as little as possible, we sought to specifically alter the affinity of this domain for chromatin. Based on the molecular model of the PWWP-domain of HDGF in complex with DNA (Lukasik et al., 2006), we predicted that positively charged residues K56 and R74 are most likely interacting with the phosphates of the host DNA (Figure 1.8 and Materials and Methods). Next, we constructed and purified two single mutants of LEDGF/p75, K56D and R74D, and the double mutant K56D-R74D and corroborated the correct overall folding with circular dichroism spectroscopy (Table 4.2).

Table 4.2 **Secondary structure of PWWP-mutants is not different from wild-type LEDGF/p75.** Circular dichroism spectroscopy on purified wild-type and mutants of LEDGF/p75. Shown is the calculated secondary structure composition of the proteins. CD spectra were measured between 195 and 260 nm on a spectropolarimeter (Jasco Benelux bv, de Meern, The Netherlands).

	WT	K56D	R74D	K56D R74D
Helix	12.80%	12.90%	12.00%	14.40%
Antiparallel	20.80%	20.20%	22.00%	17.30%
Parallel	4.40%	4.40%	4.40%	4.30%
Beta-Turn	26.30%	26.30%	25.60%	27.60%
Random Coil	36.30%	36.60%	36.60%	37.10%
Total Sum	100.50%	100.40%	100.50%	100.70%

Next, eGFP-fusions of these proteins were transiently expressed in HeLa cells. The proteins were characterized by a diffuse nuclear localisation instead of the typical heterogeneous nuclear distribution of the wild type protein (Figure 4.11 A-C), suggesting the chromatin binding properties were affected. Indeed, the photobleaching of the mutants decreased considerably (Figure 4.11 D, red curve) and the ACFs showed faster protein dynamics of eGFP-LEDGF/p75 carrying these mutations: the R74D mutation had a more pronounced effect than the K56D mutation (Figure 4.11 E, resp. red and black ACF) while the double mutant did not show an additive effect compared with the R74D single mutation (Figure 4.11 E, green ACF). The PWWP-domain of LEDGF/p75 thus ensures strong affinity chromatin binding. Importantly, the difference in cellular dynamics between the K56D and R74D mutants could not be inferred from a differential distribution of the proteins. It was only evidenced by our FCS measurements, demonstrating the sensitivity of this technique. We next verified whether the mutations in the PWWP domain of LEDGF/p75 affected chromatin tethering of IN. In FRAP experiments mRFP-IN co-expression affected the dynamics of eGFP-LEDGF/p75 K56D-R74D much less than for wild-type eGFP-LEDGF/p75 (Figure 4.11 F, red curve). In fact, the immobile fraction disappeared completely and the time scale of the mutant protein complex dynamics was much alike free eGFP-LEDGF/p75. To verify that eGFP-LEDGF K56D-R74D and mRFP-IN still interact, we used cellular fluorescence cross-correlation spectroscopy (FCCS). Indeed, marginal photobleaching and a strong interaction was observed (Figure 5.8 and Table 5.2).

Moreover, in comparison with wild-type LEDGF/p75, PWWP-mutations targeted the protein complex away from the chromatin (Figure 4.11 F). We thus conclude that the PWWP domain in LEDGF/p75 is crucial for tight chromatin tethering of IN but not for the interaction of IN with LEDGF/p75.

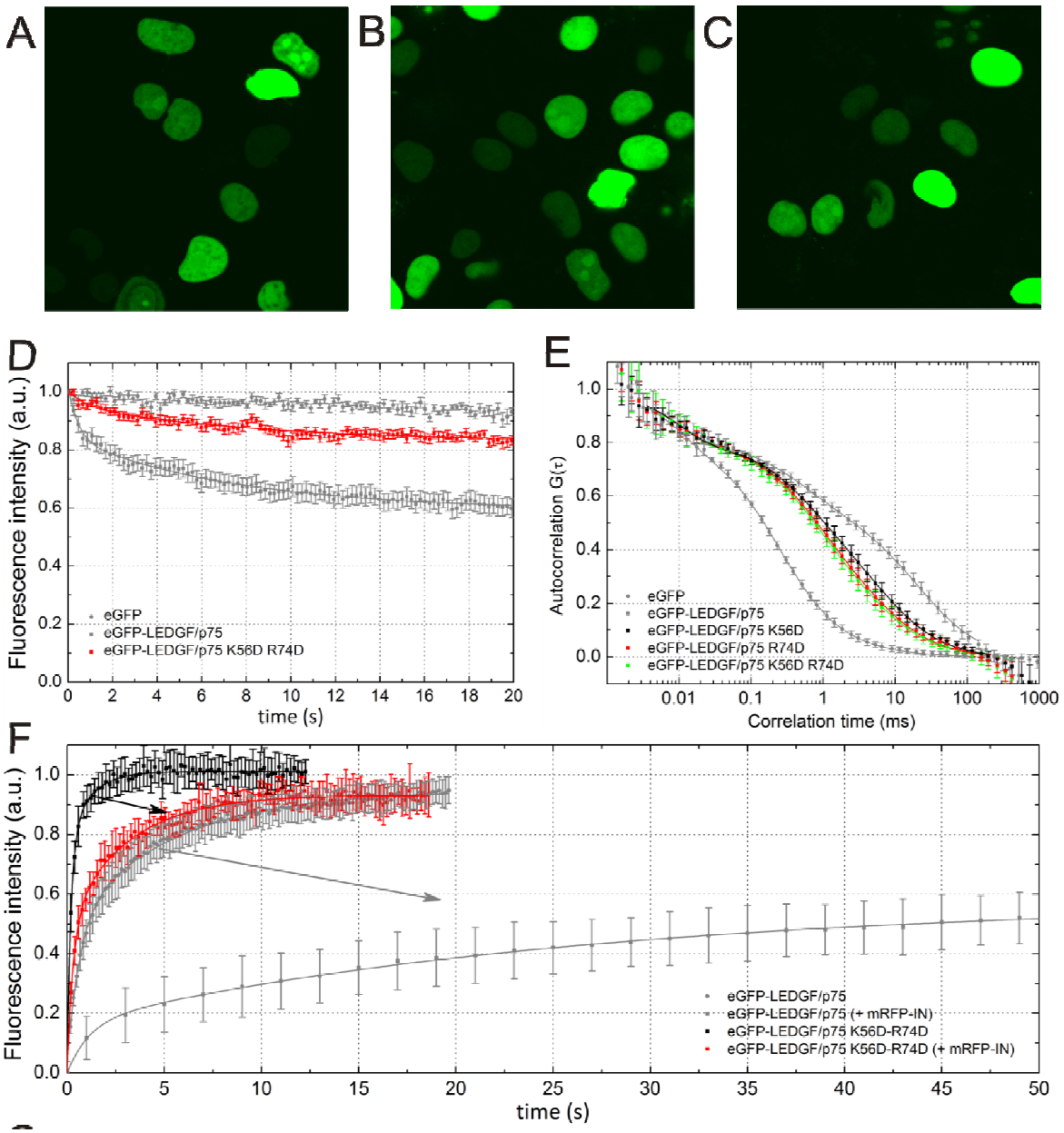

Figure 4.11 **The PWWP domain of LEDGF/p75 contributes to strong affinity chromatin binding and is crucial for chromatin tethering of HIV-1 integrase** (A-C) Confocal fluorescence images of HeLa cells expressing (A) eGFP-LEDGF/p75 K56D, (B) R74D and (C) K56D-R74D. (D) Cellular point fluorescence intensity and (E) ACF of eGFP-LEDGF/p75 K56D-R74D. (F) Cellular FRAP measurement of eGFP-LEDGF/p75 K56D-R74D in HeLa cells (n = 20). The arrows represent the shift in protein dynamics upon mRFP-IN co-expression. FRAP curves for wild-type LEDGF/p75 are plotted in gray as a reference.

4.4. Discussion

4.4.1. Fast genome scanning of LEDGF/p75 relates to its cellular function

Both Sprague et al. and independently also Phair et al. published methods in 2004 to analyse Fluorescence Recovery After Photobleaching in the presence of chromatin binding (Sprague et al., 2004; Phair et al., 2004). Recently, Michelman-Ribeiro and co-workers extended this theory for the analysis of Fluorescence Correlation Spectroscopy in the presence of chromatin binding (Michelman-Ribeiro et al., 2009). Both techniques allow observing the mobility of proteins over a defined distance, and since FCS does this over a much smaller distance than FRAP, individual chromatin association events are likely observed rather than mere diffusion of the protein. Consequently however, if many chromatin association/dissociation events occur during an individual observation, the individual events will be obscured as seemingly slow diffusion. Our TFFCS method allowed us to show for the first time that transcription factor chromatin binding kinetics can indeed be even faster than the FCS diffusion time scale and are obscured in seemingly slow 'effective' diffusion. For LEDGF/p75, a $k^*_{on} > 1.9x10^3\ s^{-1}$ and $k_{off} > 38.6\ s^{-1}$ was found, considerably faster than reported before for other transcription factors (Phair et al., 2004; Mueller et al., 2008; Michelman-Ribeiro et al., 2009). These fast kinetics, which could represent either 3D-hopping (Figure 4.6 D), 1D-sliding (Figure 4.6 E) or a combination of both (Figure 5.10), imply that LEDGF/p75 is predominantly associating with chromatin in a collision controlled fashion and finding a specific binding site amenable to tight interaction with LEDGF/p75 is thus a matter of chance. This result provides a novel insight in the field of transcription and urge a re-evaluation of former data. To that extent, we could already show with FCS that RNA polymerase II and Heterochromatin Protein 1 show slow dynamics in FCS, which are most likely due to fast chromatin binding kinetics. Specifically for LEDGF/p75, both the documented specific activation of stress responsive genes by LEDGF/p75 (Singh et al., 2001; Singh et al., 2006) and the targeting of HIV-integration to actively transcribed, yet non-specific genes (Ciuffi et al., 2005; Marshall et al., 2007; Shun et al., 2007), do fit well in the proposed model.

4.4.2. Important role of the PWWP-domain in chromatin binding of LEDGF/p75

It remained crucial to verify whether the dynamic interactions of LEDGF/p75 with chromatin are based on direct binding with DNA, as suggested before (Singh et al., 2001), or on indirect interactions with other chromatin-bound proteins. The conserved PWWP-domain of LEDGF/p75 contains a recently described general protein fold. The PWWP domain is present in proteins that carry chromatin-binding motifs and has been proposed to bind to DNA in several independent structural biology studies (Qiu et al., 2002a; Lukasik et al., 2006). Other studies however suggested, albeit without experimental evidence so far, that the PWWP-fold is implicated in non-DNA chromatin interactions (Stec et al., 2000; Maurer-Stroh et al., 2003). Although eGFP-PWWP did not show any DNA binding *in vitro*, it was shown to be localised on mitotic chromosomes, much alike full length LEDGF/p75 (Turlure et al., 2006). We studied the role of the PWWP domain of LEDGF/p75 in the intranuclear dynamic behaviour of the protein by introducing the K56D and R74D mutations in PWWP. Both mutations were designed to impede DNA binding. We showed a 4-fold increase in the overall protein dynamics and hence reduced DNA binding, as judged from the experimental ACFs, as compared to wild-type LEDGF/p75 (Figure 4.11 E). Notably, the mutant protein exhibited diffusion-dependent anomalous diffusion. This could result from interactions with the chromatin, as has been simulated before (Wedemeier et al., 2008), but could also imply that interactions with other cellular cofactors determine the dynamics when LEDGF/p75 is targeted away from the chromatin, resulting in a macromolecule suffering from obstructed diffusion (Weiss et al., 2004). Further research will address this issue. Our findings implicate that the PWWP domain is involved in direct interaction with DNA and validate the important, but possibly not unique role for PWWP in chromatin binding. Since the PWWP-domain strongly contributes to the molecular arrest and interference with the PWWP domain drastically handicaps LEDGF/p75-mediated chromatin tethering of IN, this domain may be a promising novel target for blocking LEDGF/p75 malfunction in AIDS but also oncogenesis.

4.5. Conclusion

In this chapter we have applied FCS to investigate the dynamics of LEDGF/p75 in real-time in living cells. After pinpointing the overall slow dynamics to chromatin binding, we proved that the underlying mechanism of the dynamics of LEDGF/p75 is an extraordinary fast interaction with chromatin, with multiple association-dissociation events taking place while the protein traverses the confocal measurement spot. LEDGF/p75 thus seems to mainly interact non-specifically with chromatin, although tight chromatin binding was occasionally observed. Association with HIV-1 integrase results in a segregation between an immobile and a significantly slower fraction. By decreasing the affinity of the PWWP domain of LEDGF/p75 for chromatin, we observed a moderate increase of the dynamics of LEDGF/p75 but a strong increase on the dynamics of LEDGF/p75-IN. Future work will focus on investigating the importance of this tight PWWP-domain-controlled chromatin interaction for HIV replication.

Chapter 5. Measuring the interaction of HIV-1 integrase and LEDGF/p75 in living cells

Parts of this chapter are published in:
De Rijck J., Vandekerckhove,L., Gijsbers,R., Hombrouck,A., **Hendrix,J.**, Vercammen,J., Engelborghs,Y., Christ,F., and Debyser,Z. (2006a). Overexpression of the lens epithelium-derived growth factor/p75 integrase binding domain inhibits human immunodeficiency virus replication. J. Virol. 80, 11498-11509.

Hombrouck,A., De Rijck,J., **Hendrix,J.**, Vandekerckhove,L., Voet,A., De Maeyer,M., Witvrouw,M., Engelborghs,Y., Christ,F., Gijsbers,R., and Debyser,Z. (2007). Virus Evolution Reveals an Exclusive Role for LEDGF/p75 in Chromosomal Tethering of HIV. PLoS. Pathog. 3, e47.

Hendrix, J., De Rijck, J., Gijsbers, R., Voet, A., Hotta, J., McNeely, M., Vanstreels, E., Daelemans, D., Hofkens, J., Debyser, Z., and Engelborghs, Y. Dynamic chromatin scanning of transcriptional co-activator LEDGF/p75 is arrested by HIV-1 integrase. (submitted for publication).

5.1. Introduction

After LEDGF/p75 was identified in a co-immunoprecipitation experiment together with HIV-1 integrase (IN), both proteins were shown to co-localize in interphase and mitotic cells (Cherepanov et al., 2003). The nuclear localisation of IN was shown to be dependent on the concentration of LEDGF/p75 (Maertens et al., 2003). *In vitro* pull-down experiments

confirmed that the two proteins indeed interacted specifically. Furthermore, *in vitro* LEDGF/p75 increased DNA binding of IN, and this effect was shown to be lentivirus specific (Busschots et al., 2005). Finally, it was demonstrated with fluorescence correlation spectroscopy, that the nuclear dynamics of eGFP-tagged IN were dependent on the presence of endogenous LEDGF/p75, which is an indirect proof for their interaction (Maertens et al., 2005). An assay for quantifying the direct intracellular interaction of IN and LEDGF/p75 did, however, not yet exist at the time. Fluorescence cross-correlation spectroscopy is a good candidate for such an assay, since it is one of the few techniques that can probe an interaction in a non-invasive manner in the natural environment of the cell. If such an assay could be set up, potential inhibitors of the interaction between IN and LEDGF/p75 could be validated for their activity on the IN:LEDGF/p75 interaction directly in living cell culture. In this chapter the efforts that were made to create such an assay are presented. Together with the results from 0, we finally build a first mechanistic model that unifies paradoxical notions of transcriptional co-activation and HIV-1 targeting by LEDGF/p75.

5.2. Materials and Methods

The protocols for plasmid cloning, cell culture and transfections are given in Chapter 4.

5.2.1. Fluorescence cross-correlation spectroscopy

For the FCCS measurements a commercial FCS/FCCS microscope (LSM510/ConfoCor2, Carl Zeiss, Jena, Germany) was used. Practically, enhanced green fluorescent protein (eGFP) (Heim et al., 1995) and monomeric red fluorescent protein (mRFP1) (Campbell et al., 2002) were used as fluorescent tags. The 488-nm line of the Ar^+-laser (acousto-optical tunable filter (AOTF) 0.1%, ~0.25 μW) was used to excite eGFP and the 543-nm line of the HeNe laser (AOTF 7%, ~8.6 μW) was used to excite mRFP1. The excitation light was reflected by a dichroic mirror (HFT 488/543) and focused through a C-Apochromat 40×/1.2 W Korr/0.13-0.17 objective. The fluorescence emission light was split by a second dichroic mirror (NFT 570) into two separate beam paths and passed through a 505-530 nm bandpass filter and 70-μm pinhole for eGFP fluorescence and a 600-650 nm bandpass filter and 78-μm pinhole for mRFP1 fluorescence. In each cell cytoplasm or nucleus, 10 consecutive FCCS measurements of 20 seconds were performed.

5.2.2. Qualitative FCCS analysis

For qualitative FCCS analysis, one can normalize the cross-correlation function (CCF) to the green autocorrelation function (ACF) (Saito et al., 2004; Kohl et al., 2005).

$$rel.CC = \frac{G_{CC}(0)}{G_{g,ACF}(0)} \approx fraction\ bound \qquad \text{Equation 5.1}$$

This so-called 'relative cross-correlation' parameter is 0 for no interaction and 1 for 100% interaction, when the green-to-red concentration ($N_{green,total}/N_{red,total}$) ratio is 1. Due to unavoidable spectral cross-talk and non-ideal volume overlap however, the actual CC_{rel} varied between ~0.2 (very weak or no interaction) and ~0.7 (very strong interaction). Comparable protein concentrations and concentration ratios were used to be able to compare different cellular measurements and different proteins. In living cells the expression ratio of green-to-red protein can be easily controlled by the ratio of green-to-red plasmid. This method of representing the data allows for an easy assessment of the affinity but is however not very quantitative.

5.2.3. Quantitative FCCS analysis: 'apparent' equilibrium dissociation constant

ACFs and CCFs were fitted with Equation 2.20. The amplitude of the correlation functions can be used to calculate the actual concentrations. First, the correlation functions are corrected for spectral cross-talk (section 2.3.4.2), to obtain the real observed particle numbers:

$$N_{g,total} = N_{g,ACF} = N_g + N_{gr} \qquad \text{Equation 5.2}$$

$$N_{r,total} = \frac{N_{r,ACF} - 2QN_{g,total} + \sqrt{(2QN_{g,total} - N_{r,ACF})^2 - 4QN_{g,total}(QN_{g,total} - N_{r,ACF})}}{2} \qquad \text{Equation 5.3}$$

$$N_{gr} = \frac{N_{g,total}(N_{r,total} + QN_{g,total})}{N_{CCF}} - QN_{g,total} \qquad \text{Equation 5.4}$$

Where $N_{g,total}$, N_g, $N_{r,total}$, N_r and N_{gr} are resp. the green-total, green-only, red-total, red-only and green-and-red particle numbers, $N_{g,ACF}$, $N_{r,ACF}$ and N_{CCF} are the inverses the respective correlation functions, and Q is the brightness of eGFP relative to the brightness of mRFP in the red detection channel. For the eGFP-LEDGF/p75 (or mutants thereof) and mRFP-IN couple, Q≈0.1. With these particle numbers known, the actual concentrations can be calculated:

$$[X] = \frac{N_x}{V_x N_A} \qquad \text{Equation 5.5}$$

With N_x the particle number of species X measured in volume X and N_A Avogadro's number. If protein G, tagged with eGFP, binds to protein R, tagged with mRFP, to form GR, then the affinity of this interaction defines the dissociation constant K_{diss}:

$$G + R \overset{K_{diss}}{\longleftrightarrow} GR \text{ with } K_{diss} = \frac{[G][R]}{[GR]} \qquad \text{Equation 5.6}$$

$$\text{and } [G] = [G]_0 - [GR]$$

$$[R] = [R]_0 - [GR]$$

with $[G]_0$ and $[R]_0$ the total concentrations of protein G and protein R. Importantly, to be able to accurately quantify an equilibrium with FCCS, the CCF amplitude at a given ratio of $\frac{[GR]}{[G]_0[R]_0}$ should still be well measurable above background. As a rule of thumb, to assure a measurable fraction of bound species, protein concentrations in the range K_{diss} are best used, making FCCS especially suitable for the quantification of rather strong affinity protein-protein interactions (Bacia et al., 2006)(Figure 5.1).

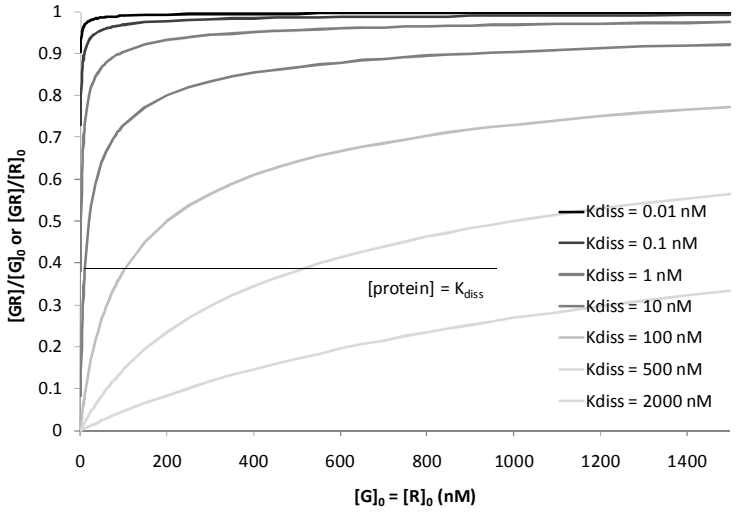

Figure 5.1 **Simulation of the percentage of complex formation in function of the total protein concentration, for a given affinity of two proteins** – Whereas at strong affinity (low K_{diss}) FCCS can already measure the interaction at nM concentrations, a significant concentration of proteins is needed if the proteins interact with weaker affinity.

5.2.4. FLIM-FRET

The fluorescence lifetime of a fluorochrome, in this case eGFP, corresponds to the relaxation time of the excited state depopulation by fluorescence. If a second fluorescent molecule with proper photophysical properties (mRFP) is at a distance smaller than ~10 nm from the first

molecule, e.g. due to protein-protein interactions, the excitation can be transferred without radiation in a process called Förster resonance energy transfer (FRET). This transfer will lower the average fluorescence lifetime of the first molecule, since its excited state will be depopulated faster. Cellular time-resolved fluorescence lifetime imaging based FRET (FLIM-FRET) can be used to study protein-protein interactions inside living cells by careful monitoring of the fluorescence lifetime at each cellular position. Because a whole cell or cell compartment rather than a diffraction limited spot is probed during one measurement, this technique is less prone to photobleaching than cellular FCCS. A commercial confocal laser scanning microscope (Leica TCS SP5, Leica Microsystems CMS GmbH, Mannheim, Germany) equipped with a pulsed laser (MaiTai-HP Ti:Sapphire laser, Newport Spectra-Physics) and time-tagged time-resolved (TTTR) time-correlated single-photon counting (TCSPC) hard/software (Picoquant GmbH, Berlin, Germany) was used for FLIM-FRET. The 880-nm laser light was focused through a HCX-PL-APO-63x/NA-1.2/water objective (Leica) and eGFP emission was captured through a 535/50 bandpass filter. Practically, the LSM was adjusted to detect very low concentrations of expressed protein by single-photon excitation. A cell with a suitable expression ratio of green-to-red protein (< 1:5) was magnified and two-photon excitation of eGFP at 1.5% of the 80-Mhz 880-nm pulsed laser output allowed performing TCSPC-based FLIM-FRET. The fluorescence decay curve was calculated for the whole cytoplasm or nucleus and fitting was performed by a mono- or biexponential reconvolution fitting procedure. The instrument response function was determined by TCSPC of a solution of malachite green under the same experimental conditions. All FLIM-FRET experiments were carried out at the K.U.Leuven Rega Institute for Medical Research.

5.3. Results

5.3.1. LEDGF/p75 and IN form an immobile intracellular complex

As a fluorochrome couple for these experiments, we used the best fluorescent protein couple that was available at the time, enhanced green fluorescent protein eGFP (Heim et al., 1995) and the first true monomeric red fluorescent proteins, mRFP1 (Campbell et al., 2002), which have a good spectral emission separation (eGFP $\lambda_{em,max}$=507 nm, mRFP $\lambda_{em,max}$=607 nm) and have a brightness that is compatible with FCCS measurements. We generated two fusion constructs, eGFP-LEDGF/p75 and mRFP-IN, co-expressed these hybrids in HeLa cells and

performed an FCCS experiment at an homogenous region in the nucleoplasm (Figure 5.2 top). During the first seconds of the measurements strong photobleaching occurred (Figure 5.2 A). As a consequence of this, correct auto- and cross-correlation of the fluorescence signals was no longer possible (Figure 5.2 B).

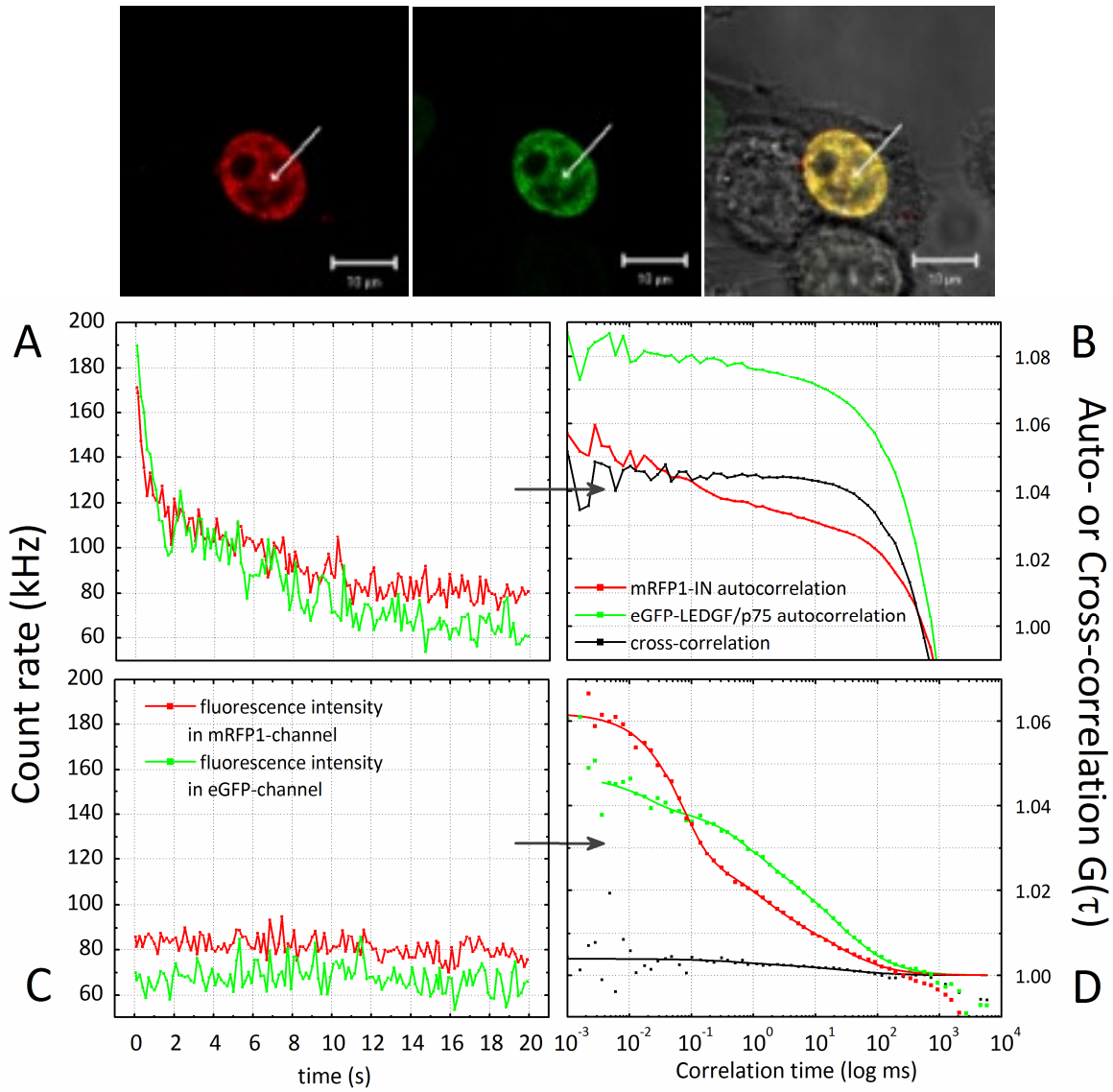

Figure 5.2 **FCCS measurements of wild-type eGFP-LEDGF/p75 are prohibited due to photobleaching**. (top) Confocal LSM image of HeLa cells expressing mRFP-IN (red) and eGFP-LEDGF/p75 (green) and the overlay of the two images (yellow). (bottom)(A-B) In cells co-expressing wild-type eGFP-LEDGF/p75 and mRFP-IN strong photobleaching occurred during the first 20 seconds of the measurement resulting in erroneous correlation curves. (C-D) After the signal reached a plateau, no significant cross-correlation was observed.

The complex of these proteins is likely immobile, resulting in the photobleaching. Once the signals had reached a plateau (Figure 5.2 C), correct auto- and cross-correlation curves could be calculated and analysed, but the lack of a significant cross-correlation curve amplitude, which is a signature an interaction of the two proteins, lead us to conclude that no significant interaction between LEDGF/p75 and IN could be observed in the nucleoplasm (Figure 5.2 D).

To pinpoint this immobilisation to an interaction of IN and LEDGF/p75, we measured the dynamics of eGFP-tagged IN under different experimental conditions. First, we expressed eGFP-IN in wild-type HeLa cells and measured the fluorescence intensity versus time at a certain region in the nucleoplasm. We observed variable photobleaching from cell to cell (Figure 5.3 top left).

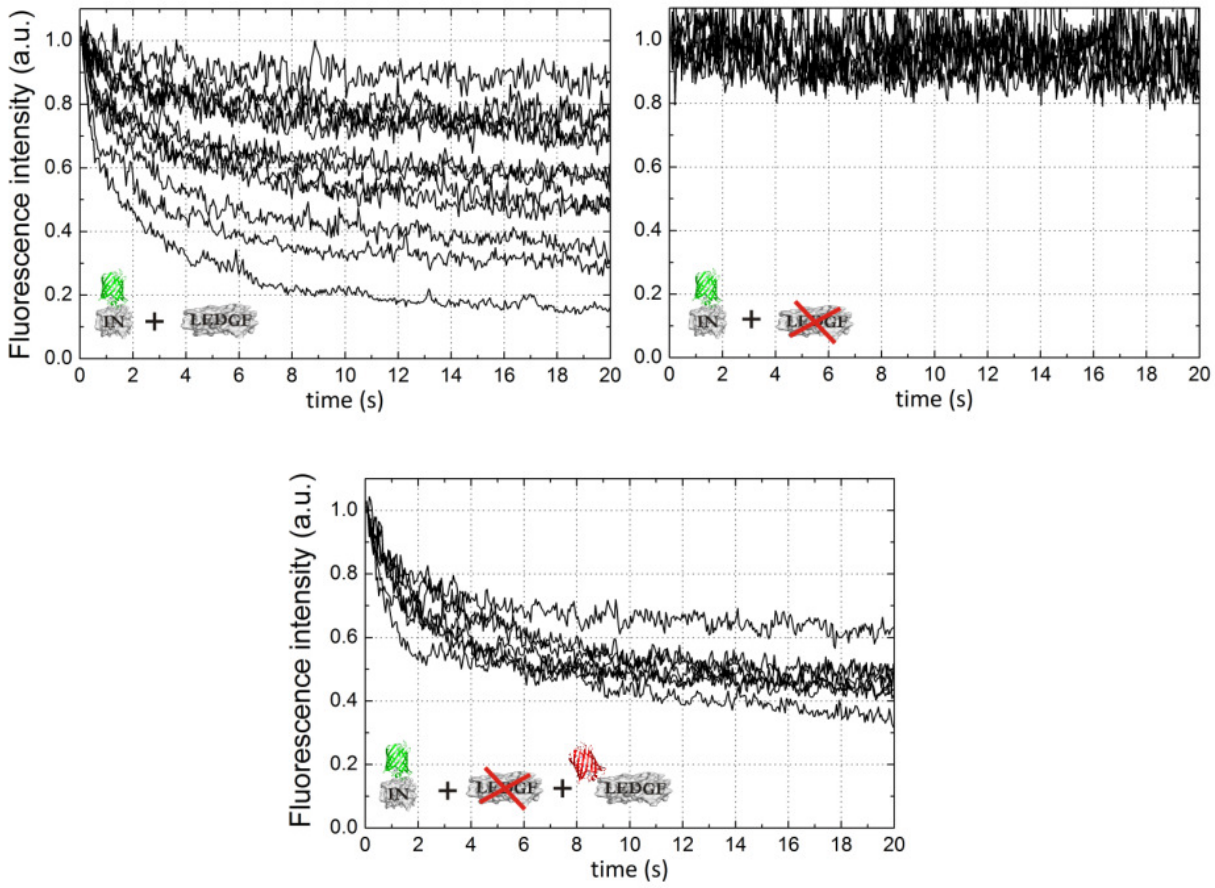

Figure 5.3 **The dynamics of IN are dictated by LEDGF/p75** – Point fluorescence measurements of eGFP-IN in (top left) wild-type HeLa cells, (top right) HeLa-p75KD cells and (bottom) HeLa-p75KD cells back-complemented with 500 nM mRFP-LEDGF/p75.

Then, we performed the same experiment in HeLa cells with a stable knockdown of LEDGF/p75 (HeLa-p75KD). No significant photobleaching of eGFP-IN could be observed in HeLa-p75KD cells (Figure 5.3 top right). This implies that the differential photobleaching of IN in normal HeLa cells was due to an interaction with LEDGF/p75 and corroborates that IN behaves as a free protein in the absence of LEDGF/p75 (Maertens et al., 2005). In a third experiment, we back-complemented HeLa-p75KD with RNAi-resistant mRFP-LEDGF/p75. When measuring eGFP-IN in cells expressing at least 500 nM of mRFP-LEDGF/p75 (Figure 5.3 bottom), a large fraction of eGFP-IN was photobleached, proving that photobleaching of IN was indeed dependent on the concentration of LEDGF/p75. These result confirms the observations of Busschots et al., who showed *in vitro* that LEDGF/p75 increases the affinity of IN for DNA considerably (Busschots et al., 2005).

In conclusion, FCCS experiments on cell co-expressing eGFP-LEDGF/p75 and mRFP-IN were not successful due to protein immobilization. We clearly showed that LEDGF/p75 is the determinant for immobilization. In light of the known chromatin binding properties of LEDGF/p75 (Chapter 4), this immobilization is most likely due to chromatin binding. In the following section, we will investigate whether oligomerization or aggregation of IN contributes to the immobilization.

5.3.2. IN-IN interactions as probed with FLIM-FRET

It is known that the IN binding domain (IBD) of LEDGF/p75 needs at least an IN catalytic core domain dimer to bind to (Figure 1.9)(Cherepanov et al., 2005a). For our intracellular assay to probe the interaction of IN and LEDGF/p75, it is thus crucial that IN can still oligomerize. So in a next step, we determined whether fluorescent protein tagged IN monomers could interact with each other in a cellular context to form a binding pocket for LEDGF/p75, in the absence of the latter. We did this by using fluorescence lifetime imaging based Förster resonance energy transfer (FLIM-FRET). We co-expressed mRFP- and eGFP-tagged IN in a >5:1 ratio in HeLa-p75KD cells. At this concentration ratio, each eGFP-IN will associate with at least one mRFP-IN if IN oligomerizes. When eGFP and mRFP in such a complex are within Förster distance apart, measuring the fluorescence lifetime of eGFP allows to probe the interaction. Indeed, both in the cytoplasm and in the nucleus the fluorescence lifetime of eGFP-IN decreased in the presence of mRFP-IN (Table 5.1 part 1), indicating a close intracellular interaction, independent of LEDGF/p75 or chromatin. This decrease in fluorescence lifetime was not as pronounced as for our control protein mRFP-

eGFP, with eGFP as a control (Table 5.1 part 1). Since FRET depends on the intermolecular distance, the distance between the fluorophores in eGFP-IN complexed with mRFP-IN is likely larger. Since the interaction was detected at low concentrations of IN, as judged from the confocal fluorescence image, it classifies as a strong affinity interaction ($K_{diss} \leq 100$ nM). Of concern, fluorescence correlation spectroscopy measurements on cellular eGFP-IN did not reveal any significant change in molecular brightness with respect to eGFP, making it impossible to quantify this IN-IN interaction with FCS (data not shown).

Table 5.1 **HIV-1 integrase and LEDGF/p75 interactions as quantified with FLIM-FRET** – Subcellular fluorescence decays were fitted by a single exponential reconvolution procedure, as described in the experimental procedures. s.d. = standard deviation, n = number of independent cellular measurements. Overexpression of protein 1 and 2, as well as back-complemented LEDGF/p75 was achieved via transient transfection of plasmid DNA in the HeLa cell line specified. LEDGF↓: LEDGF/p75 stable knockdown cell line. LEDGF/BC↑ and LEDGF/BC K56D-R74D↑: in the LEDGF/p75 stable knockdown cell line back-complementation with respectively wild-type and K56D-R74D mutant LEDGF/p75 was achieved after transfection with a plasmid carrying silent mutations in the LEDGF/p75 gene as not to be targeted by the miRNA.

	protein 1	protein 2	Hela type	Back-comp.	location	n	lifetime (ns)	±	s.d.
1	eGFP		HeLa		whole cell	11	2.43	±	0.02
	mRFP-eGFP		HeLa		whole cell	10	1.95	±	0.05
	eGFP-IN		LEDGF↓		nucleus	40	2.37	±	0.05
	eGFP-IN		LEDGF↓		cytoplasm	20	2.35	±	0.04
	eGFP-IN	mRFP-IN	LEDGF↓		nucleus	42	2.19	±	0.02
	eGFP-IN	mRFP-IN	LEDGF↓		cytoplasm	20	2.15	±	0.04
2	eGFP-LEDGF		LEDGF↓		nucleus	30	2.35	±	0.03
	eGFP-LEDGF	mRFP	LEDGF↓		nucleus	10	2.32	±	0.03
	eGFP-LEDGF	mRFP-LEDGF	LEDGF↓		nucleus	30	2.32	±	0.05
3	eGFP-IN	mRFP-LEDGF	LEDGF↓		nucleus	5	2.20	±	0.02
	eGFP-LEDGF	mRFP-IN	LEDGF↓		nucleus	10	2.23	±	0.03
4	eGFP-IN		LEDGF↓	LEDGF/BC↑	nucleus	10	2.24	±	0.02
	eGFP-IN	mRFP	LEDGF↓	LEDGF/BC↑	nucleus	10	2.24	±	0.03
	eGFP-IN	mRFP-IN	LEDGF↓	LEDGF/BC↑	nucleus	20	2.06	±	0.04
5				LEDGF/BC					
	eGFP-IN		LEDGF↓	K56D-R74D↑	nucleus	10	2.37	±	0.02
				LEDGF/BC					
	eGFP-IN	mRFP-IN	LEDGF↓	K56D-R74D↑	nucleus	10	2.15	±	0.02
6	eGFP-IN	mRFP-IN	HeLa	LEDGF K150A↑	cytoplasm	10	2.19	±	0.01

Likewise, we determined whether LEDGF/p75 monomers interact in the living cell (Table 5.1 part 2). However, the fluorescence lifetime of eGFP-LEDGF/p75 was not altered upon mRFP-LEDGF/p75 co-expression, suggesting that dimerization is not occurring. This is in

accordance with *in vitro* dimerization studies of the PWWP-domain of human HDGF, that showed that, in absence of heparin, the dimerization affinity was very weak (K_{diss} = 14.4 mM) (Sue et al., 2007).

Next, we verified if the intracellular IN-IN interaction in absence of LEDGF/p75 could be an aggregation instead of an oligomerization. Aggregation is accompanied by the formation of large complexes containing multiple fluorochromes. *In vitro* purified mRFP-IN induces such aggregation of purified eGFP-LEDGF/p75, resulting in the appearance of bright fluorescence spikes in both red and green fluorescence traces (Figure 5.4). The method of Van Craenenbroeck et al. was used to analyse these burst-containing fluorescence traces (Van Craenenbroek et al., 2001), but quantitative analysis was not possible due to problems with *in vitro* protein stability. Since we never observed momentary bursts of fluorescence in cells, a well known indication for aggregating protein complexes, we conclude that the IN:IN complex likely represents a dimer or tetramer, but not a higher oligomer. Next, we investigated whether the co-expression of LEDGF/p75 influenced IN:IN interaction. We compared the fluorescence lifetime of eGFP-IN in cells depleted for endogenous LEDGF/p75 with the same cells back-complemented with LEDGF/p75. Surprisingly, the fluorescence lifetime of eGFP-IN decreased from 2.37 ns to 2.24 ns in the presence of LEDGF/p75 (Table 5.1 part 3). Possibly, the immobilisation we described in the previous section alters the physicochemical environment of the eGFP tag affecting the fluorescence lifetime. Co-expression of mRFP-IN further decreased the lifetime, but not to a significantly higher extent than in cells without LEDGF/p75.

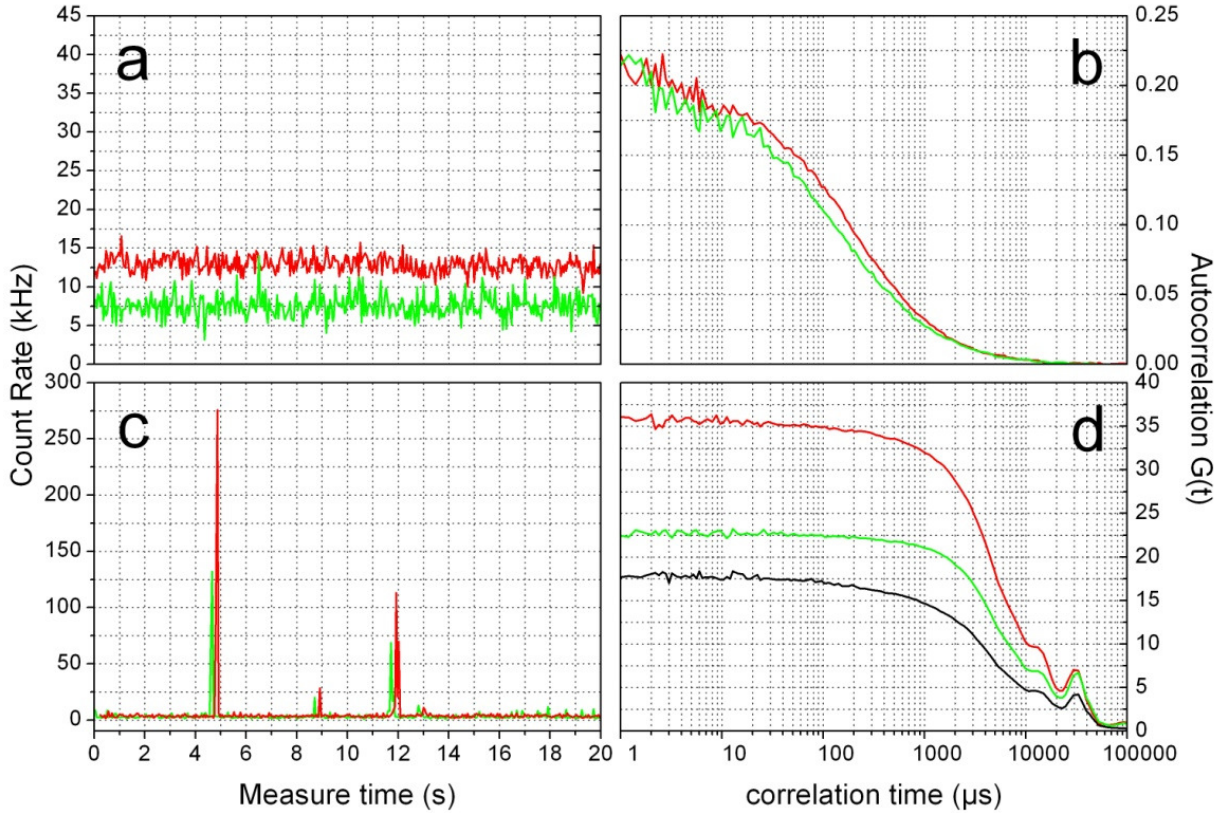

Figure 5.4 **FCS and FCCS analysis of *in vitro* purified mRFP-IN and eGFP-LEDGF/p75** – (a-b) FCS measurement on a sample containing only mRFP-IN or only eGFP-LEDGF/p75 in a 20 mM Hepes pH 7.4 buffer with 150 mM NaCl. Both proteins show a stable fluorescence signal and artifact-free ACF. Results are plotted on the same graph. (c-d) FCCS measurement on a mixture of both proteins in the same buffer. Bright fluorescence spikes appear, with erronous FCCS curves as a result.

In conclusion, individual IN monomers, tagged with a fluorescent protein, do interact in cells to form dimers or possibly tetramers, but not higher oligomers. LEDGF/p75 monomers do not interact in the living cell nucleus. Under conditions of LEDGF/p75 mediated immobilization of IN, we cannot detect an alteration of the IN:IN interaction using FLIM-FRET.

5.3.3. Strategies to cancel out the immobility of the complex

FCCS is a fluorescence fluctuation technique. If a protein complex is immobile, fluorescence doesn't fluctuate and no interaction is observed. To still be able to study the interaction between LEDGF/p75 and IN in the living cell with FCCS, we saught for methods to inhibit the immobilization. Different strategies were followed, based on the protein structure of LEDGF/p75 (Figure 1.8); Deletion of the complete N-terminal p52-domain of

LEDGF/p75, mutation of the nuclear localisation signal in LEDGF/p75 or mutation of the chromatin binding PWWP domain.

5.3.3.1. Deletion of the complete N-terminal p52-domain of LEDGF/p75

The N-terminal part of LEDGF/p75, that comprises the major part of its alternative splice variant p52 (Singh et al., 2000a), contains nearly all chromatin binding domains. Deletion of only this p52 part from LEDGF/p75 (Δ325) should completely inhibit chromatin interactions of LEDGF/p75. We constructed and expressed eGFP-Δ325 in HeLa cells, together with mRFP-IN. eGFP-Δ325 still exhibited residual nuclear localization, likely due to two *helix-turn-helix* motifs in the Δ325-part of LEDGF/p75 (Shinohara et al., 2002) (Figure 5.5). More importantly and quantitatively, FCS measurements showed that more than 89% of the protein population exhibited free diffusion, in comparison to only 35% for wild-type LEDGF/p75 (Table 4.1). Surprisingly, mRFP-IN does no longer localize in the nucleus in cells expressing eGFP-Δ325 (Figure 5.5). This cytoplasmatic retention implies the formation of a protein complex that is too big to enter the nucleus by passive diffusion and imposes a strong block on HIV replication (De Rijck et al., 2006). When we performed FCCS measurements in the cytoplasm of such cells, no photobleaching occured and a high relative cross-correlation was observed (Equation 5.1), rel.CC = 0.62, pointing to a specific interaction of the two proteins (Figure 5.5 and Table 5.2). The measured relative cross-correlation amplitude of 0.62 would imply that 62% of eGFP-Δ325 is present in a complex with mRFP-IN. As a control we also performed an FCCS experiment with mRFP instead of mRFP-IN. Clearly, the cross-correlation was considerably lower, rel.CC = 0.15.

We verified the specificity of the FCCS analysis by introducing the D366A point mutation in Δ325. From *in vitro* binding studies it is known that the interaction with IN is decreased by this mutation and from crystallography it is known that the D366 residue lies exactly in the loop that extends from LEDGF/p75 into the cleft formed by the IN-CCD dimer (Figure 1.11). The CCF amplitude is lowered to the background level, rel.CC = 0.15, by introducing this mutation in eGFP-Δ325, validating that FCCS is sensitive to probe the interaction. Moreover, the diffusion time of eGFP-Δ325 D366A in cells expressing mRFP-IN is significantly faster than that of eGFP-Δ325 under the same conditions, which is also an indirect proof that the interaction is lower.

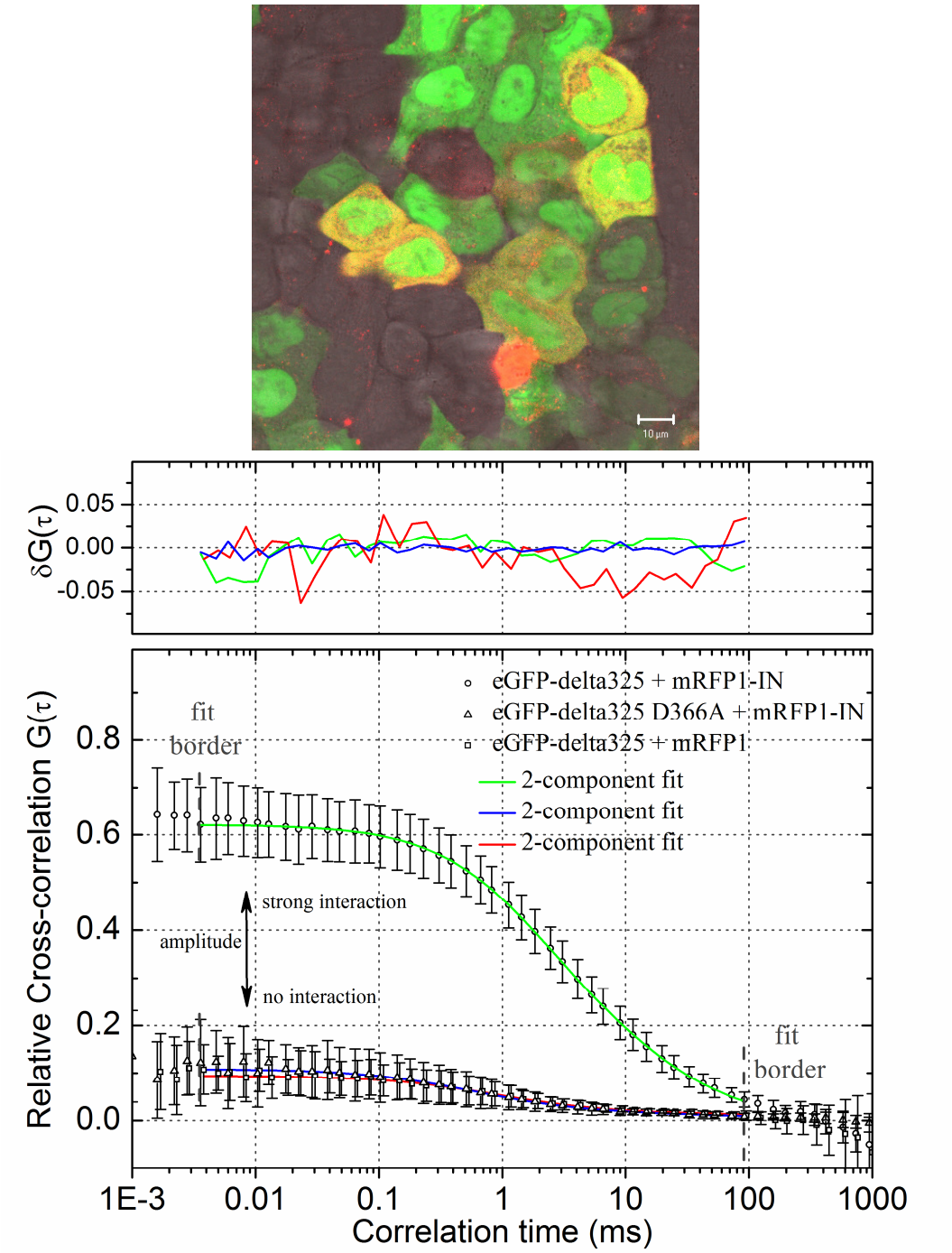

Figure 5.5 **The eGFP-Δ325/mRFP-IN cytoplasmic FCCS assay – (top panel)** Confocal images of HeLa cells transiently expressing eGFP-Δ325 and mRFP-IN. **(bottom panel)** FCCS measurements on mRFP-IN and eGFP-Δ325. Colors are explained in the legend.

In conclusion of these measurements, when the p52 chromatin binding part of LEDGF/p75 is removed, a cytoplasmic complex with IN is formed. Importantly, this complex is no longer immobile and FCCS becomes a sensitive tool to probe the interaction.

Table 5.2 **Interaction of mRFP-IN with LEDGF/p75 variants measured with FCCS –** The relative cross-correlation was calculated with Equation 5.1. $[G]_0$ = the total concentration of green protein, n = number of independent cellular measurements, s.d. = standard deviation, n.p. = fit not possible, a.u. = arbitrary units, * = negative control where mRFP1 instead of mRFP1-IN was used.

LEDGF variant	properties	n	$G_c(0)/G_g(0)$±s.d.	$N_{g,total}/N_{r,total}$±s.d.	$[G]_0$ ± s.d. (nM)
Cytoplasm					
eGFP-Δ325		22	0.62 ± 0.08	1.23 ± 0.27	647 ± 281
eGFP-Δ325 + mRFP1*		16	0.15 ± 0.08	1.04 ± 0.50	1025 ± 625
eGFP-Δ325 D366A	mIBD	15	0.15 ± 0.08	0.69 ± 0.36	298 ± 259
eGFP-LEDGF/p75 K150A	mNLS	33	0.50 ± 0.11	0.98 ± 0.27	498 ± 243
Nucleus					
eGFP-LEDGF/p75	wild type	n.p.			
eGFP-LEDGF/p75 R74D	mPWWP	10	0.48 ± 0.08	1.19 ± 0.26	1701 ± 777
eGFP-p75 K56D-R74D	mPWWP	19	0.50 ± 0.06	1.02 ± 0.20	1516 ± 512

5.3.3.2. Mutation of the NLS of LEDGF/p75

It has been shown that LEDGF/p75 carries a canonical nuclear localization signal, ^{148}GRKRKAEKG^{156} and that a single amino acid mutation, K150A, is sufficient to disrupt the natural localization of LEDGF/p75 (Maertens et al., 2004; Vanegas et al., 2005). We constructed eGFP-LEDGF/p75 K150A to probe the interaction of full-length LEDGF/p75 with mRFP-IN in the cytoplasm. First, we verified the cellular localization of the protein. Within 24 hours after transfection, the localization of both proteins is cytoplasmic (Figure 5.6 top left). After mitosis however, LEDGF/p75 K150A is retargeted back to the nucleus because of chromatin binding (Vanegas et al., 2005) (Figure 5.6 top right). Next, with fluorescence correlation spectroscopy we checked if the mutated protein, in spite of the amino acid mutation, retained nuclear dynamics similar to LEDGF/p75 (Figure 5.6 bottom). Indeed, in the nucleus the dynamics of eGFP-LEDGF/p75 K150A was identical to wild-type LEDGF/p75. In the cytoplasm faster dynamics were logically observed. Moreover, the ACF could only be described well with a model for anomalous diffusion. This suggests LEDGF/p75 interacts with other cellular components in the cytoplasm, resulting in a protein complex that experiences obstructed diffusion.

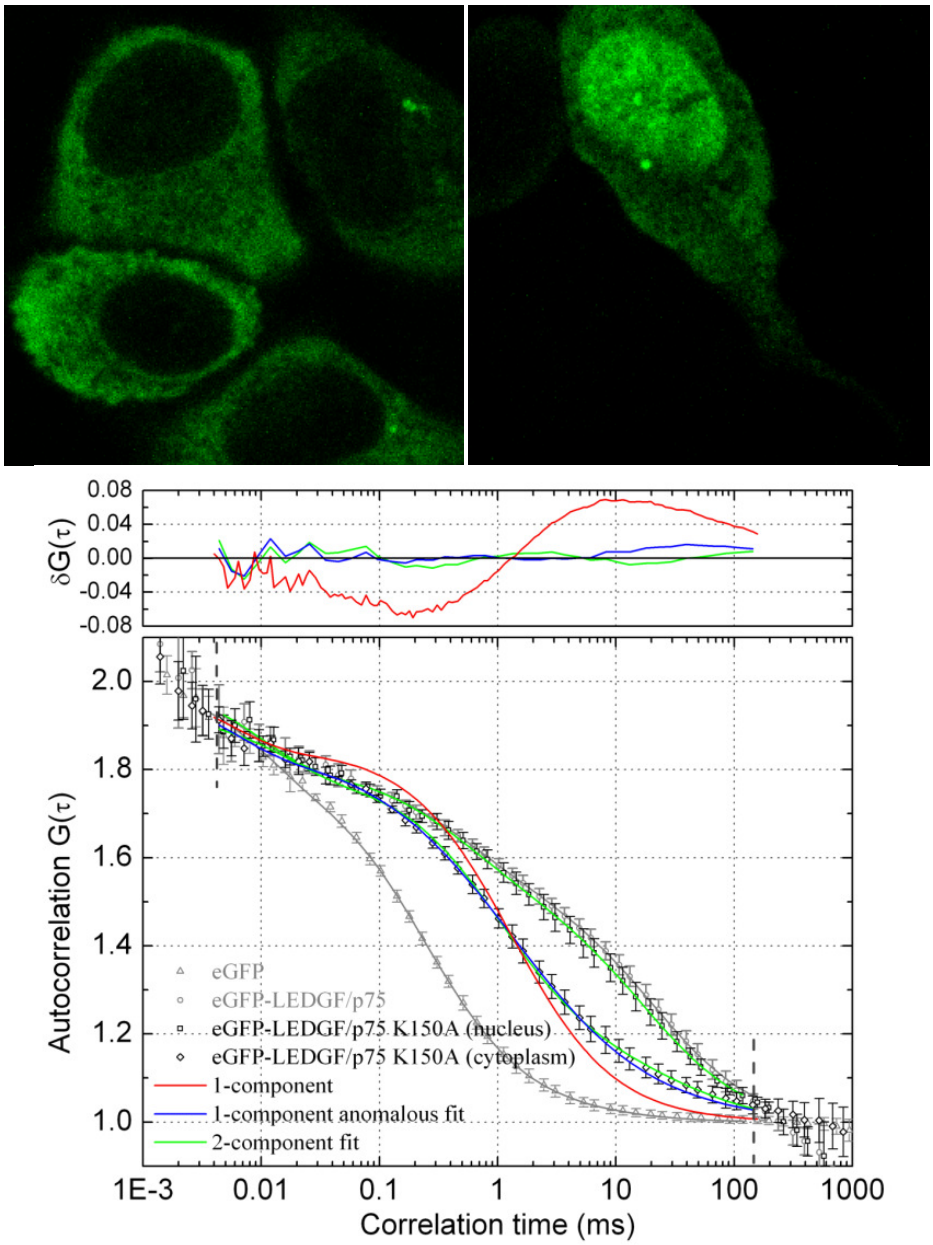

Figure 5.6 **LEDGF/p75 K150A and IN interact in the cytoplasm** - **(top panels)** Cellular distribution of eGFP-LEDGF/p75 K150A (top left) before and (top right) after mitosis. **(bottom panel)** Normalized cellular autocorrelation function of eGFP-LEDGF/p75 K150A in the cytoplasm and nucleus of HeLa cells. Fitted curves and residuals are color coded in the figure legend. ACFs for eGFP and eGFP-LEDGF/p75 are shown for reference.

Then, we co-expressed mRFP-IN. As for eGFP-Δ325, mRFP-IN was retained in the cytoplasm by eGFP-LEDGF/p75 K150A (Figure 5.7 top). FCCS showed again a strong interaction could again be observed (Figure 5.7 bottom) which was within experimental error, the same as for mRFP-IN + eGFP- Δ325. The dynamics of this complex were significantly

slower than for the eGFP-Δ325/mRFP-IN complex, as judged by the diffusion part of the CCF, suggesting a bigger protein complex is formed.

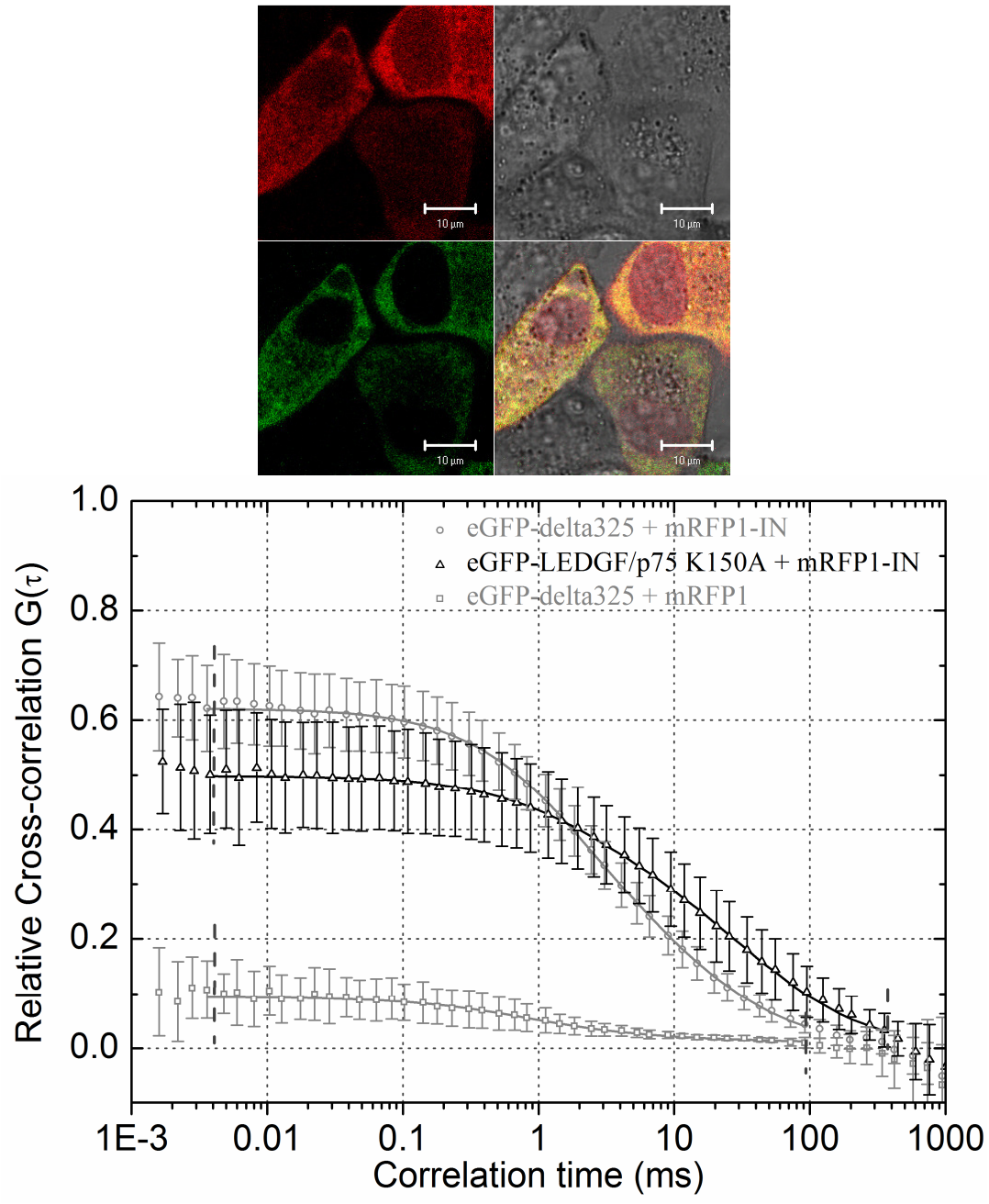

Figure 5.7 **The eGFP-LEDGF/p75/K150A-mRFP-IN FCCS assay – (top panel)** Intracellular localisation of eGFP-LEDGF/p75 K150A (green) and mRFP-IN (red) in HeLa cells. **(bottom panel)** Normalized cellular cross-correlation function of eGFP-LEDGF/p75 K150A and mRFP-IN in the cytoplasm. Solid line: fit with 2-component model. CCFs for eGFP-Δ325 and mRFP-IN or mRFP are shown as a reference (De Rijck et al., 2006).

5.3.3.3. Mutation of the chromatin binding PWWP domain of LEDGF/p75

Endogenous LEDGF/p75 is a nuclear protein, it thus would be useful if we could also measure the interaction with IN in the nucleus. As we have discussed in detail in Chapter 4, the N-terminal chromatin binding PWWP domain in LEDGF/p75 can be mutated at residues K56 and R74, to generate a LEDGF/p75 variant that retains nuclear localization, but does no longer form a tightly chromatin bound complex upon co-expression of IN. We generated eGFP-fusions of the LEDGF/p75 K56D-R74D mutant and expressed this protein together with IN and performed FCCS measurements. Nuclear localization was observed (Figure 5.8 top) and again a high relative cross-correlation was observed (Figure 5.8 bottom and Table 5.2). In summary, by using the PWWP-mutant of LEDGF/p75, FCCS can be used to probe the interaction with IN also in the nucleus.

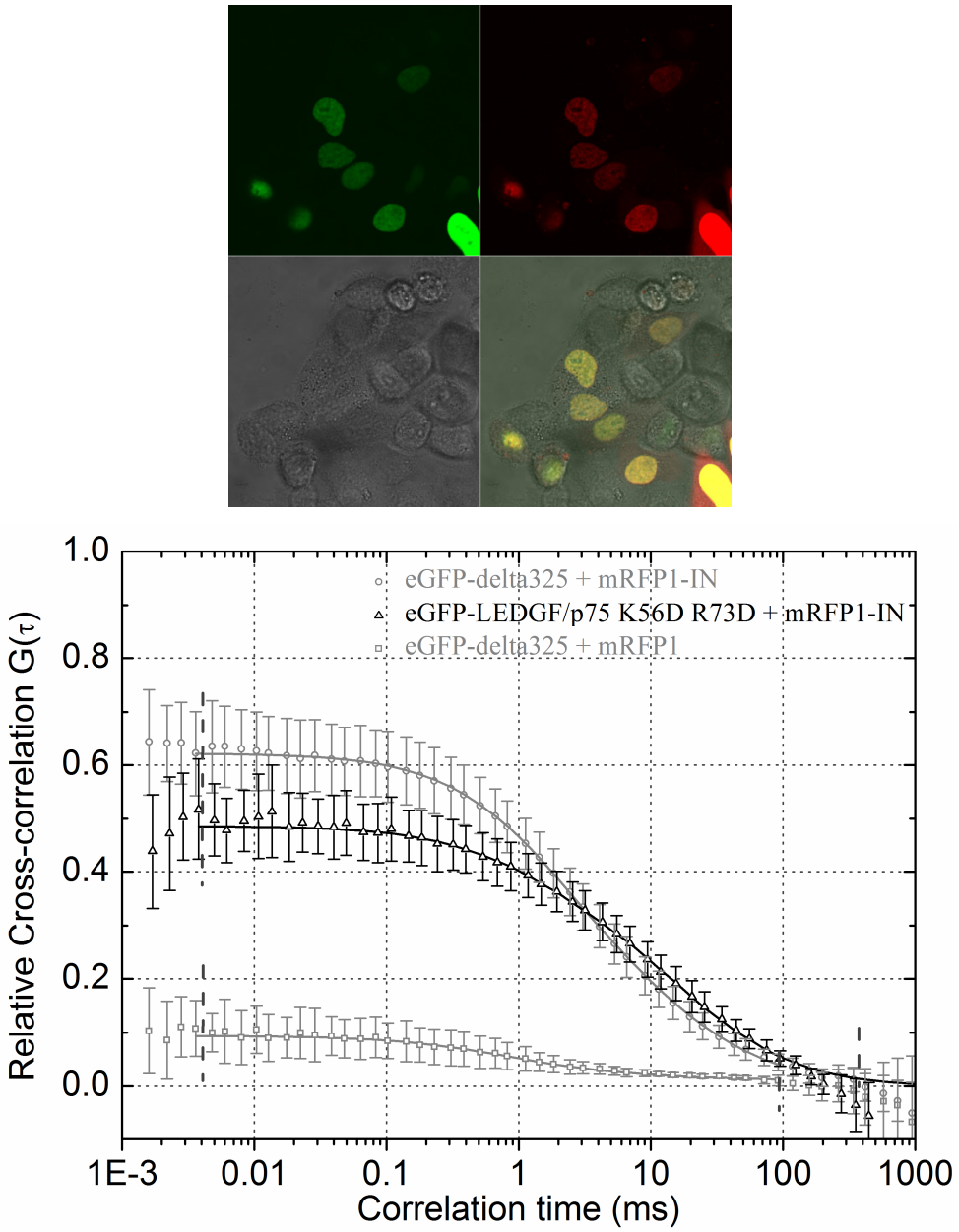

Figure 5.8 **The nuclear eGFP-LEDGF/p75 K56DR74D-mRFP-IN FCCS assay – (top panel)** Intracellular localisation of eGFP-LEDGF/p75 K56DR74D (green), mRFP-IN (red) and overlay, together with a transmission image overlay. **(bottom panel)** FCCS measurements in cells expressing eGFP-LEDGF/p75 K56D-R74D and mRFP-IN (n = 20). The high amplitude of the black CCF means a strong protein-protein interaction. Error bars = s.d.

We also tested both the NLS-mutant and PWWP-mutant LEDGF/p75 with FLIM-FRET, to see if the IN-IN interaction was affected by the presence of full-length LEDGF/p75, but in absence of chromatin binding. First we co-expressed in our LEDGF/p75 knockdown cell line eGFP-IN/mRFP-IN (<1:5 plasmid ratio) and untagged LEDGF/p75 K56D-R74D (IN:LEDGF plasmid ratio 1:1) for which we know that chromatin tethering is impaired but

not the interaction with IN. This resulted in high expression of mutant LEDGF/p75 in nearly all cells that expressed the fluorescent IN, as confirmed with immunocytochemistry. Additionally fluorescent IN, randomly distributed under LEDGF/p75 knockdown conditions, was retargeted to the nucleus upon LEDGF/p75 K56D-R74D co-expression. The fluorescence lifetime decrease of eGFP-IN in cells co-expressing mRFP-IN but no LEDGF/p75 equalled the decrease in cells co-expressing LEDGF/p75 K56D-R74D (Table 5.1 part 5), suggesting that no alteration of the IN-IN interaction occurs in cells expressing full-length chromatin-binding-deficient LEDGF/p75. Secondly we co-expressed untagged LEDGF/p75 K150A, with a dysfunctional NLS in HeLa cells. Again no significant change in FRET was detected (Table 5.1 part 6), suggesting that also in the cytoplasm, LEDGF/p75 has no detectable effect on the IN-IN interaction. In summary, in the absence of chromatin tethering we find no evidence that LEDGF/p75 affects the IN-IN interaction in cells.

5.3.4. Quantitative FCCS analysis

As we have shown in the previous sections, the FCCS technique can be used to provide conclusive information on the presence or absence of an intracellular protein-protein interaction, provided that the dynamics of the protein complex are not too slow. With respect to LEDGF/p75 and IN, in section 5.3.3 we measured the interaction of mRFP-IN with different LEDGF/p75 variants in cells. From Figure 5.7 and Figure 5.8, it might seem that the affinity of eGFP-LEDGF/p75-K150A and eGFP-LEDGF/p75-K56DR74D for mRFP-IN is lower than that of eGFP-Δ325. The ultimate quantification of affinity would be, however, an intracellular binding constant. As stated in the Materials and Methods section, by correcting the correlation functions for cross-talk, this becomes possible. Importantly, FRET can have a large influence on the correlation functions too (as we will see in Chapter 6), so we verified the FRET contribution. As can be seen in Table 5.1, the minor decrease of the fluorescence lifetime of eGFP-IN from 2.24 ns in cells expressing untagged LEDGF/p75 to 2.20 ns in cells expressing mRFP-LEDGF/p75, suggests a neglegible contribution of FRET. We re-analysed the FCCS measurements and calculated the equilibrium binding constant K_{diss} (Figure 5.9 and Table 5.3). As expected, the affinity of Δ325 and LEDGF/p75-K150A for IN in the cytoplasm is similar. We calculate a binding constant for dissociation of around 400 nM, suggesting a fairly strong affinity.

For the PWWP-mutant LEDGF/p75 K56D-R74D we observed a weaker apparent affinity for IN (K_{diss} ~900 nM). This result is logical since these experiments were not performed in

cells with a stable knockdown of LEDGF/p75 (the knockdown cell line was not available at the time). Thus, in the nucleus there will be a vast concentration of endogenous LEDGF/p75 present which competes with the fluorescent LEDGF/p75 for binding to mRFP-IN. Secondly, it is known that other co-factors can bind to LEDGF/p75, at the integrase binding domain (IBD) (see section 1.6.3), and these co-factors also localize only in the nucleus. In other words, there will be either off-titration of fluorescent LEDGF/p75 by endogenous LEDGF/p75, or off-titration of fluorescent IN by cellular co-factors in the nucleus. In both cases, the apparent K_{diss} for LEDGF/p75-IN in the nucleus will be lower than the 'real' K_{diss}, which is what we observed. Next, for our negative control, eGFP-Δ325/mRFP, the binding constant was artificially high (K_{diss} = ~49 μM), which is logical since these proteins cannot bind to each other. Finally, for the D366A mutant of Δ325, we could still calculate a binding constant (K_{diss} = ~3400 nM), suggesting that at least some interaction with IN is maintained.

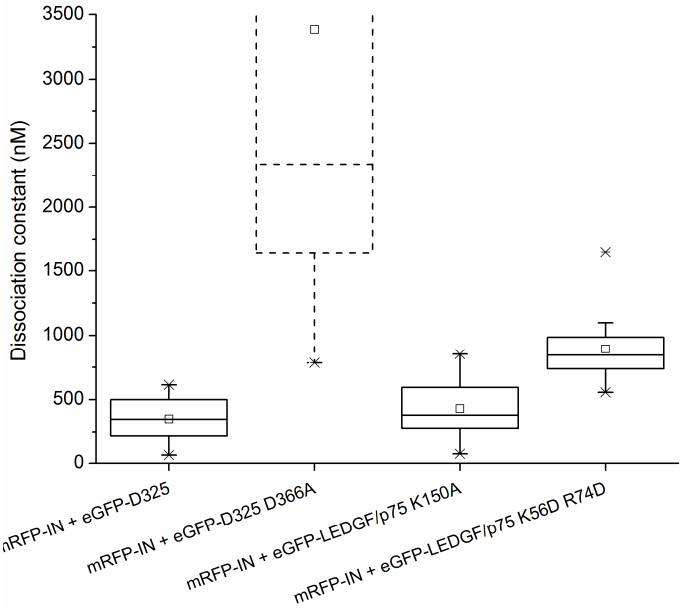

Figure 5.9 **Box-plot representation of the calculated K_{diss} from the FCCS measurements of IN with different LEDGF/p75 variants** – The box limits represent the 25% and 75% percentile, the whiskers represent the 5% and 95% percentile, stars represent outliers, the line in the box is the median and the square in the box is the mean.

In conclusion, when FCCS measurement are corrected for cross-talk, we can calculate intracellular binding constants of IN with LEDGF/p75 (and variants thereof). Quantitative analysis of this interaction is useful, because *in vitro* identified inhibitors of this interaction can now be quantitatively tested inside cells.

Table 5.3 **Quantification of the interaction between mRFP-IN and eGFP-LEDGF/p75 variants –** The K_{diss} calculation is explained in de materials and methods section, For the calculation of K_{diss}, the outliers were not included. n = number of independent measurements, s.e.m. = standard error of mean, * = negative control where mRFP1 instead of mRFP1-IN was used.

LEDGF variant	properties	n	K_{diss} ± s.e.m.	
Cytoplasm				
eGFP-Δ325		29	350 ± 31	nM
eGFP-Δ325 + mRFP1*		14	49 ± 12	**µM**
eGFP-Δ325 D366A	mIBD	17	3380 ± 690	nM
eGFP-LEDGF/p75 K150A	mNLS	28	429 ± 38	nM
Nucleus				
eGFP-p75 K56D-R74D	mPWWP	16	890 ± 65	nM

5.4. Discussion

5.4.1. Fluorescent protein tagging does not affect chromatin tethering of IN

We wanted to verify if the observed chromatin tethering is a consequence of the tag on IN. After transient expression, eGFP-IN displayed a diffuse nuclear localisation, with residual fluorescence in the cytoplasm ((De Rijck et al., 2006) and data not shown) as has been reported before for eGFP-IN (Maertens et al., 2005), GST-IN (Gallay et al., 1997) and IN (Cherepanov et al., 2000). When transiently co-expressing mRFP-LEDGF/p75, eGFP-IN localized exclusively in the nucleus with a much more heterogenous distribution pattern overlapping with that of mRFP-LEDGF/p75 (data not shown). This pattern has been reported before for unlabelled IN/LEDGF/p75 (Cherepanov et al., 2003) and HcRed1-IN/eGFP-LEDGF/p75 (Maertens et al., 2003) and is consistent with both proteins interacting with each other.

It should be noted that the same pattern was observed for cells co-expressing IN-3xFLAG, IN tagged with a triple FLAG tag for immunocytochemistry, and eGFP-LEDGF/p75 (data not shown). Furthermore, our measurements clearly showed that eGFP-IN nor mRFP-IN associate with chromatin in the absence of LEDGF/p75 (resp. Figure 5.3 and data not shown), as had been suggested before for untagged IN (Maertens et al., 2003). Of concern,

untagged IN has been suggested to form tetramers in the cell (Cherepanov et al., 2003). According to our FLIM-FRET measurements (Table 5.1) tagged IN monomers do interact so that oligomerization still can occur. Furthermore, the marked decrease in the dynamics of eGFP-LEDGF/p75 (Figure 4.9) is observed after co-expression of both mRFP-IN and IN-3xFLAG (data not shown). In conclusion, although it is not known whether the stoichiometry of IN is affected by the fluorescent tag, tight chromatin association appears to be independent of its presence.

5.4.2. Higher order complex formation causes deceleration of LEDGF/p75-IN

It has been shown that LEDGF/p75 can act as a molecular tether coupling IN to the chromatin (Maertens et al., 2003; Cherepanov et al., 2003; De Rijck et al., 2006; Llano et al., 2006b; Hombrouck et al., 2007; Shun et al., 2008). Much research effort has also gone into identifying the chromatin binding domains of LEDGF/p75 (Turlure et al., 2006; Llano et al., 2006b; Shun et al., 2008) and to investigating the complex formed with IN *in vitro* (Cherepanov et al., 2003; Faure et al., 2005; Cherepanov et al., 2005b; Busschots et al., 2007; McKee et al., 2008; Michel et al., 2009). However, only few attempts have been made to define the actual intracellular complex between LEDGF/p75, IN and chromatin (Cherepanov et al., 2003; Maertens et al., 2005). The mere interaction of LEDGF/p75 and IN would logically not lead to the strong chromatin interaction we observed. Possible explanations are:

(1) *An allosteric effect*. Binding of LEDGF/p75 to IN could influence the conformation of IN in an allosteric way, increasing the affinity of the latter for chromatin. If this would be true, then only the IBD portion of LEDGF/p75 (Figure 1.8) should have the same effect. Disproving evidence is the fact that eGFP-Δ325:IN is not tightly associated to chromatin (De Rijck et al., 2006). Secondly, binding of IN to LEDGF/p75 could allosterically influence LEDGF/p75. However, since the IBD and chromatin binding sites are located in well separated domains in LEDGF/p75 (Figure 1.8), we assume this option is unlikely. Thirdly, LEDGF/p75 could change the conformation of the chromatin when binding to it. As we have clearly shown, the known chromatin binding motif PWWP in LEDGF/p75 grants tight chromatin association to the IN complex, which might be interpreted as increasing the accessibility of chromatin for IN. Further research will have to demonstrate this.

(2) *A cooperative effect*. Binding of LEDGF/p75 to IN might shift the stoichiometry of IN. Indeed, it has been shown *in vitro* that LEDGF/p75 can promote IN tetramerisation (McKee

et al., 2008; Michel et al., 2009), but it remains to be proven if this occurs as well in the living cell. With FLIM-FRET, we could clearly observe a defined oligomerisation of IN (Table 5.1), but we did not observe an effect of LEDGF/p75 on the fluorescence lifetime of eGFP-IN/mRFP-IN in the living cell, in the presence or absence of chromatin binding. Most likely, the mobile albeit slow fraction (66%) in our FRAP measurements (Figure 4.10) represents a defined stoichiometry, presumably $IN_4LEDGF/p75_2$, as has been suggested from *in vitro* work (Cherepanov et al., 2003; Faure et al., 2005; McKee et al., 2008; Michel et al., 2009). The 34% immobile fraction we observed with FRAP likely represents a higher order IN:LEDGF/p75 complex. It has been shown that recombinant IN can undergo a DNA induced polymerization *in vitro* (Vercammen et al., 2002), forming large complexes of DNA and IN. In addition, it has recently been suggested that IN in the viral particle could be present as a hetero-multimeric paracrystalline complex with multiple LEDGF/p75 proteins bound (Hare et al., 2009b). In summary, IN can oligomerize inside the cell to form a complex of defined or higher order stoichiometry and since more than one LEDGF/p75 binds to one IN complex, the dynamics shift to a much slower time scale.

(3) *A competitive effect*. Finally, cellular cofactors that reduce the chromatin binding of LEDGF/p75, could be competing with IN that by itself promotes chromatin binding of LEDGF/p75. Recently, it has been shown that JPO2, a myc-interacting protein, and PogZ, a domesticated transposase, interact specifically with LEDGF/p75 and can compete with IN for binding to LEDGF/p75 (Maertens et al., 2006; Bartholomeeusen et al., 2007; Bartholomeeusen et al., 2009). Our observation that diffusion of Δ325 in the nucleus is slower than in the cytoplasm (Table 4.1) might indeed be related to these interactions. Insight in the contribution of these and other binding partners to the dynamics of LEDGF/p75 will certainly be of value to completely understand the cellular function of LEDGF/p75.

5.4.2.1. The affinity between IN and LEDGF/p75

For LEDGF/p75 to bind to IN, IN has to be present as a dimer, as crystallographic studies has shown (Cherepanov et al., 2005a). In fact, we were able to show with FLIM-FRET that IN was always present in at least a dimeric form, irrespective of LEDGF/p75. Recently, the affinity of IN-CCD and LEDGF/p75-IBD was determined *in vitro* with Surface Plasmon Resonance to be K_{diss} = 10.9 nM. Our calculated cellular K_{diss} values of full-length mRFP-IN and eGFP-Δ325 or LEDGF/p75-K150A are around 400 nM, which would mean the affinity is weaker in cells. Considering the possibility of specific or aspecific competition in cells, this

results might not come as a surprise. Finally, it must be stressed that the technique of dual-color FCCS with fluorescent proteins, only allows to measure apparent dissociation constants, as long as a positive control did not validate the technique yet. In other words, protein binding affinities of different protein couples relative to each other are possible, but are not to be interpreted as an absolute affinity.

5.5. Conclusion

We have set up a set of intracellular assays to probe for the presence or absence of the LEDGF/p75-IN interaction, based on fluorescence cross-correlation spectroscopy. Due to a higher order stoichiometry, the interacting complex likely contain multiple LEDGF/p75 and IN partners, and we had to target the protein complex away from chromatin in order to quantify the interaction with FCCS. The biological significance of the slow dynamics of the LEDGF/p75-IN complex may be advantageous for the viral integration complex, because it confers more time to IN to perform proper integration in the host cell genome. We speculate that the mechanism of the transcriptional coactivator, i.e. aspecific genome scanning followed by arrest at defined sites of gene activation, probably through association with the transcription machinery, has been hijacked by the lentiviral integration machinery to guarantee both high frequency of integration sites and proviral gene activation after integration. In Figure 5.10, we have summarized our results on the dynamic chromatin scanning of LEDGF/p75, and how IN likely manipulates these interactions.

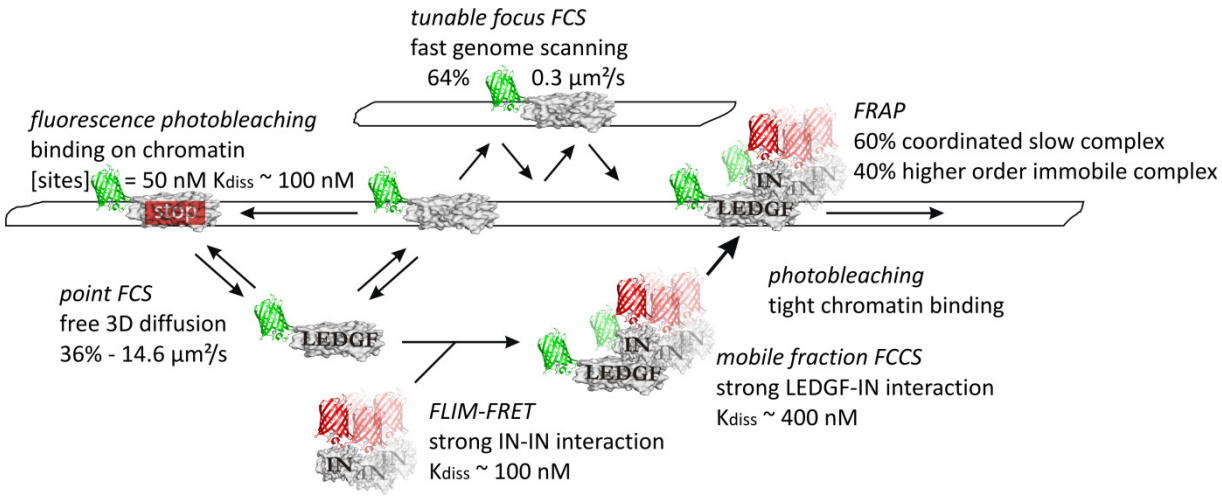

Figure 5.10 **Unifying model of the dynamic interactions of LEDGF/p75 with chromatin and of LEDGF/p75-mediated chromatin tethering of HIV-1 integrase** – LEDGF/p75 is present in an equilibrium between free nucleoplasmic diffusion and dynamic chromatin scanning. The majority of time, LEDGF/p75 is dynamically scanning the chromatin and occasionally, is halted on the chromatin by strong-affinity binding. This might explain the paradoxical fact that LEDGF/p75 has been shown to both bind DNA specifically at HSE and STRE promoters and aspecifically in transcriptionally active regions. HIV-1 integrase forms oligomers in cells and upon association with LEDGF/p75 is tethered strongly to chromatin. The PWWP domain of LEDGF/p75 has a crucial role in strong chromatin tethering, but is not required to form the LEDGF/p75-IN complex.

Chapter 6. FCCS for measuring protein-protein interactions quantitatively in cells

6.1. Introduction

Fluorescence cross-correlation spectroscopy holds great promise for becoming the biophysical tool of choice for quantifying protein-protein interactions in cells, because in one single measurement, both the absolute concentrations of the two interacting partners as the concentration of their complex can directly be measured. This makes it possible to quantify the absolute affinity of two proteins inside living cells, something no single other technique can do. Fluorescent proteins are favorable for the application of FCCS in cells for several reasons. Genetic fusion of a protein-of-interest with a fluorescent protein is straightforward and fusion constructs can be expressed simply by transfection methods or by generating a cell line with stable expression. The fluorescent protein can be placed at either the N- or C-terminus of the protein and if monomeric fluorescent proteins are used, the stoichiometry of labelling is always known beforehand. Additionally, fluorescent protein fusions can be expressed in the natural environment of the cell with all the benefits of eukaryotic expression, such as post-translational modifications and heat shock protein assisted folding. Finally, their photophysical properties can compete with those of fluorescent dyes. On the other hand, chemical *in vitro* labelling of proteins with fluorescent dyes requires *in vitro* expression, purification, labelling and introduction in cells, and one usually has little control over the labelling position or stoichiometry. In the worst case, *in vitro* purification of the protein is not possible because of protein stability issues.

Since the first reports of FCCS with fluorescent proteins (Saito et al., 2004; Baudendistel et al., 2005; Kohl et al., 2005), there have been many applications of the technique to study a variety of biological systems (for a recent review, see(Hwang and Wohland, 2007)). An important step in the development of the technique is the calculation of absolute intracellular parameters such as equilibrium binding constants, and first attempts in this direction have been made (Oyama et al., 2006; Shi et al., 2009). However, due to the complexity of the experimental setup and the often limited understanding of fluorescent protein photophysics, quantitative interpretation of experimental FCCS data can be challenging. In this chapter, we perform quantitative FCCS in living cells on a genetic fusion of the enhanced green fluorescent protein eGFP (Heim et al., 1995) with either one of two red fluorescent proteins, mRFP1, the first true monomeric RFP (Campbell et al., 2002) or mCherry, an improved point mutant of the latter (Shaner et al., 2004). We wish to investigate in detail the effects of FRET and cross-talk on the experimental ACFs and CCF. Finally, we discuss on future methods to improve intracellular FCCS.

6.2. Materials and Methods

The protocols for plasmids cloning, cell lines and transfections are given in Chapter 4. The details of the FCCS and FLIM-FRET setups are explained in detail in Chapter 5. All FLIM-FRET experiments were carried out at the K.U.Leuven Rega Institute for Medical Research under the supervision of Dr. Dirk Daelemans.

6.3. Results

6.3.1. Low cross-correlation for fluorescent fusion proteins

The fusion constructs mRFP-eGFP and mCherry-eGFP were expressed in HeLa cells and FCCS measurements were performed. Under ideal conditions, i.e. 100% confocal volume overlap and ideal fluorochrome optical properties, the eGFP and RFP autocorrelation functions (ACFs) and the cross-correlation function (CCF) should overlap completely. However, since the focal volume created by the 543-nm HeNe laser used to excite RFP is slightly bigger (factor 1.37) than the one created by the 488-nm Ar-ion laser used to excite eGFP, we expect the green ACF amplitude to be slightly higher, and the red ACF amplitude to be slightly lower than the CCF amplitude, as we have simulated in section 2.3.4. Experimentally, both for mRFP-eGFP as for mCherry-eGFP the red ACF amplitude is

highest (Figure 6.1 left panels) and moreover, the CCF amplitude is much lower than would be expected from a fusion protein. More specifically, the sharp rise of the eGFP and RFP ACFs at a sub-μs time scale is APD afterpulsing, which was discussed in detail in Chapter 3 and can be omitted from the fitting (see fit borders in Figure 6.1 right panels). It must be stated however, that at low count rates, afterpulsing can contribute to the ACF even in the 10s of microseconds time scale. Secondly, although it is not trivial to separate diffusion and photophysics from the RFP ACFs, a significant (50%) fraction of the RFP ACF represents a ~50-μs dark state. In Chapter 7 these dark states have been characterized in detail. Finally, during the first 20 seconds of the measurement, on average 20% of the RFP signal photobleached, both for mRFP1 and mCherry. At the end of the 10×20 seconds measurement sequence, up to 60% of the RFP signal had photobleached. On the other hand, for eGFP only 10% of the signal had photobleached when the measurement was finished. It is known that both mRFP1 and mCherry are much less photostable than eGFP (Shaner et al., 2008), and the laser power we used was the lowest power possible to get adequate signal/background. After fitting with the proper models, the concentration amplitudes of the correlation functions were plotted (Figure 6.1 right panels). Clearly, the RFPs have a much bigger fraction of dark states populations and the particle number seems a factor of 1.15 (mCherry) and 1.3 (mRFP) lower than the particle number for eGFP.

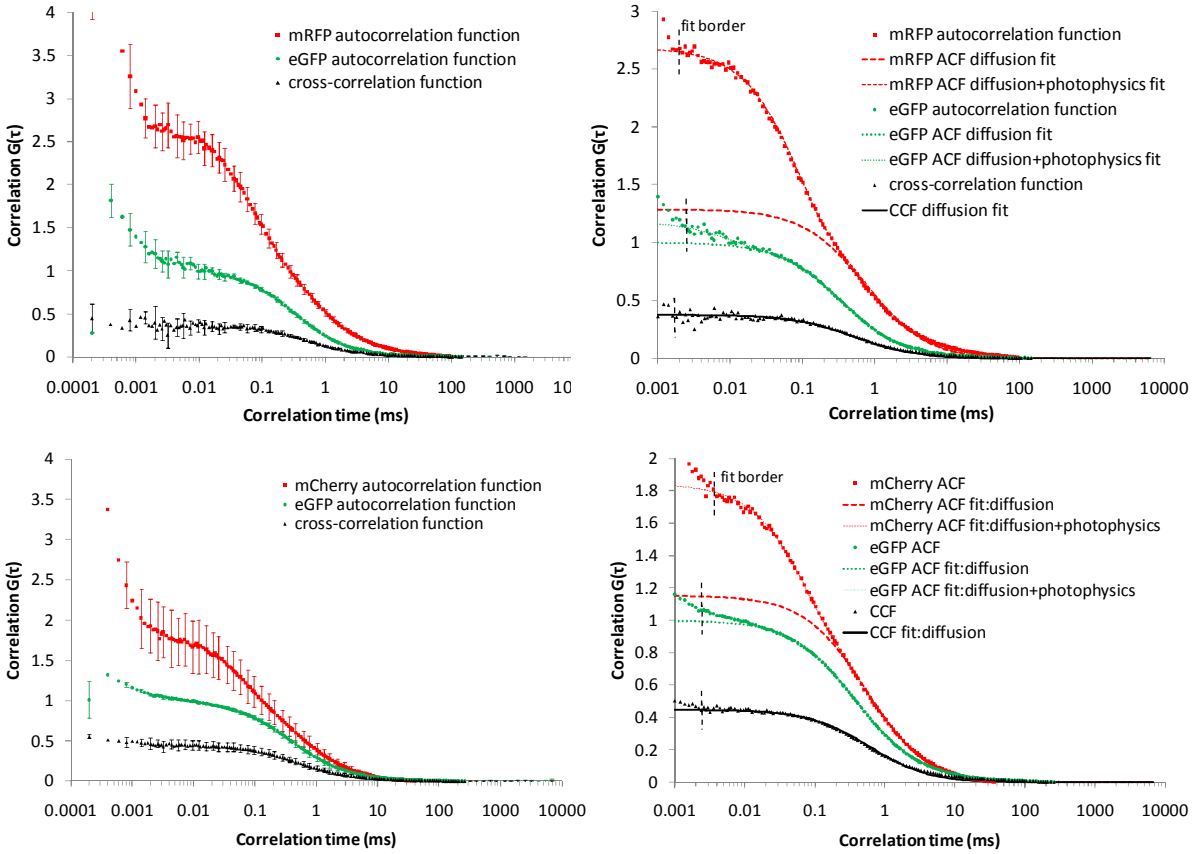

Figure 6.1 **FCCS experiment on fluorescent protein fusions** – mRFP-eGFP (top), and mCherry-eGFP (bottom). The curves represent an average of ten individual measurements in different cells. In each cell 10 successive FCCS measurements were performed and averaged. This set of 2 ACFs and 1 CCF was finally normalized by multiplying with the green particle number. Error bar = s.d.. (right) Non-linear least squares fitting with Equation 2.17. For clarity, with the obtained parameters, we plotted the ACFs with or without the photophysics, to reveal the diffusion component amplitude on the graph. Different fit lines are explained in the legend. In Chapter 7 the RFP dark states are characterized in detail.

6.3.2. 'Apparent' particle numbers

With the concentration amplitude of the ACFs we can calculate an 'apparent' particle number of green and red molecules. These are not the absolute, concentration related particle numbers, but are biased by non-idealities in the experimental setup, as we discussed in the introduction (sections 2.3.4-2.3.5). First, due to the non-ideal volume overlap, particle numbers observed in different channels correspond to different concentrations, which is straightforward to correct for if the microscope is calibrated properly. Volume overlap has been discussed in the introduction (section 2.3.4) and a proper microscope calibration method will be discussed in section 3.3.6. Secondly, green molecules contribute to the autocorrelation function in the red detection channel due to emission cross-talk, which has to

be corrected for (section 2.3.4.2). Thirdly, although RFPs and eGFP have a large spectral separation (the emission maxima are separated by 100 nm) the emission spectrum of eGFP overlaps well with excitation spectrum of the RFPs, making them good FRET partners as well, which also has to be corrected for (section 2.3.4.3). Stoichiometry does not pose an extra complexity, since the fusion proteins have a 1-1 stoichiometry.

6.3.3. Quantifying FRET with FLIM-FRET

We performed whole-cell FLIM-FRET measurements to quantify FRET in mRFP-eGFP and mCherry-eGFP. Therefore, we measured the fluorescence decay of eGFP in cells expressing these constructs, with the fluorescence decay of eGFP in cells only expressing eGFP as a control. As illustrated in Figure 6.2 for mRFP-eGFP, the fluorescence decay of eGFP in mRFP-eGFP is shifter to a faster time scale, indicative of FRET. While for eGFP-only a mono-exponential fit model described the data well (Figure 6.2, green fit line), for the fusion protein a mono-exponential model did not describe the decay well (Figure 6.2, blue fit line), suggesting the existence of multiple species in the fusion protein. When using a bi-exponential model the fit quality was much better (Figure 6.2, red fit line) and a significant fraction existed with a lifetime close to the donor-only lifetime. This fraction could be fitted as well with the slow lifetime fixed to the donor-only lifetime (2.43 ns), strongly suggesting it represents the donor without the presence of a FRET acceptor.

The results of the fitting are summarized in Table 6.1. For mCherry-eGFP 61% and for mRFP-eGFP 63% of the eGFP-moieties in the fusion protein still have a fluorescence lifetime corresponding to donor-only (resp. F_g=0.61 and F_g=0.63). In other words, a large fraction of the fusion proteins can be considered to contain a permanent or transient dark acceptor. As we have observed during the FCCS experiments, due to the limited photostability of the RFP fluorochromes, up to 60% of the fluorescence photobleached during the experiment, but these results could as well suggest that a vast percentage of the RFP fluorochromes have not fully matured.

Next, from the fluorescence lifetime of eGFP in the presence or absence of the RFP FRET acceptor, the FRET efficiency E can be calculated:

$$E = 1 - \frac{\tau_{DA}}{\tau_D} \qquad \text{Equation 6.1}$$

with τ_{DA} the fluorescence lifetime of the donor in the presence of acceptor and τ_D the lifetime in absence of acceptor. Like this, we calculated the FRET efficiency in a functional fusion

protein to be around 75% (Table 6.1). This is not unexpected, since the distance between the fluorochromes in such a fusion protein is expected to be around 40 Å, which is what we calculated.

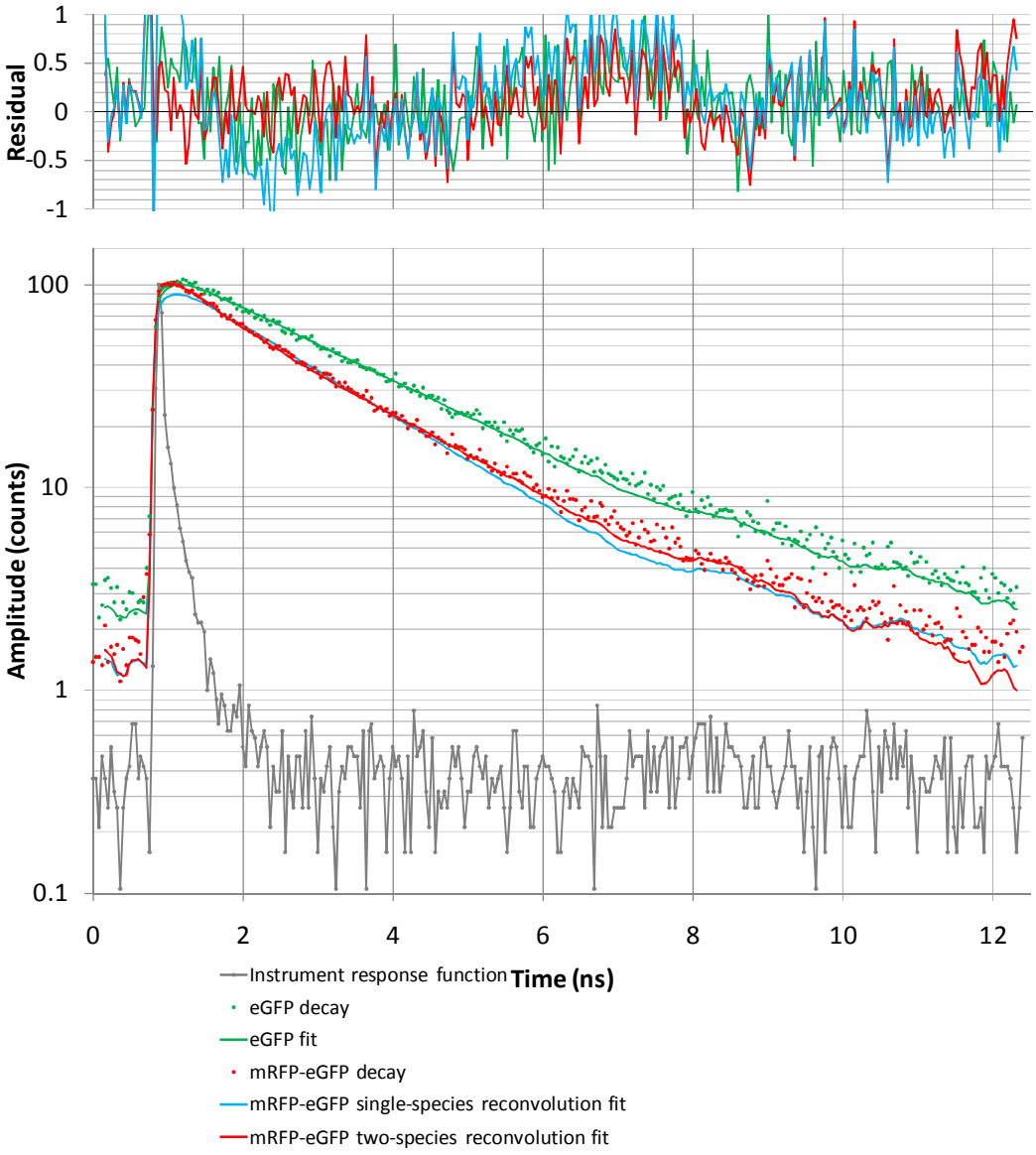

Figure 6.2 **Fluorescence decay histograms of eGFP (green dots) and eGFP in the mRFP-eGFP construct (red dots)** – Whole cell-decay of single representative cells. The instrument response function (as measured with malachite green is presented in gray. Fits and residuals are explained in the legend.

Table 6.1 **Analysis of FLIM-FRET experiments in living HeLa cells** – Reconvolution fitting was used to fit the fluorescence decays. The eGFP decay could be fitted well already with a one-compenent model, whereas the fusion proteins had to be fitted with a two-component model. E = FRET efficiency. Intermolecular distances (r) were calculated with Equation 2.46. Chi² expresses the goodness-of-fit. Parameters are given ± s.d.

protein	n	lifetime 1 (ns)	F_g	lifetime 2 (ns)	F_{gr}	E	r (Å)	chi²
eGFP	11	**2.43** ± 0.02						1.02 0.05
		2.66 ± 0.09	0.82 ± 0.04	0.83 ± 0.31	0.18 ± 0.04			0.97 ± 0.03
mRFP-eGFP	10	**1.95** 0.05						
		2.34 ± 0.07	**0.63** ± 0.02	0.59 ± 0.06	0.37 ± 0.02	**0.76** ± **0.08**	**38.9**	0.96 ± 0.03
mCherry-eGFP	10	2.30 ± 0.1	**0.61** ± 0.04	0.58 ± 0.04	0.39 ± 0.04	**0.76** ± **0.06**	**38.8**	0.95 ± 0.05

6.3.4. Correcting the eGFP autocorrelation function

With these parameters, we can now correct the eGFP ACF. The green ACF only has to be corrected for FRET with Equation 2.46:

$$N_{g,app} = N_{g,total,real}\frac{\left(F_g + (1-E)F_{gr}\right)^2}{F_g + (1-E)^2 F_{gr}} \qquad \text{Equation 6.2}$$

where the FRET efficiency E, the fraction green-only F_g and the fraction of green-and-red F_{gr} have to be measured with another technique. In Figure 6.3, we have plotted some simulated corrections of the eGFP ACF and the corrected ACF with the experimentally determined FRET efficiency.

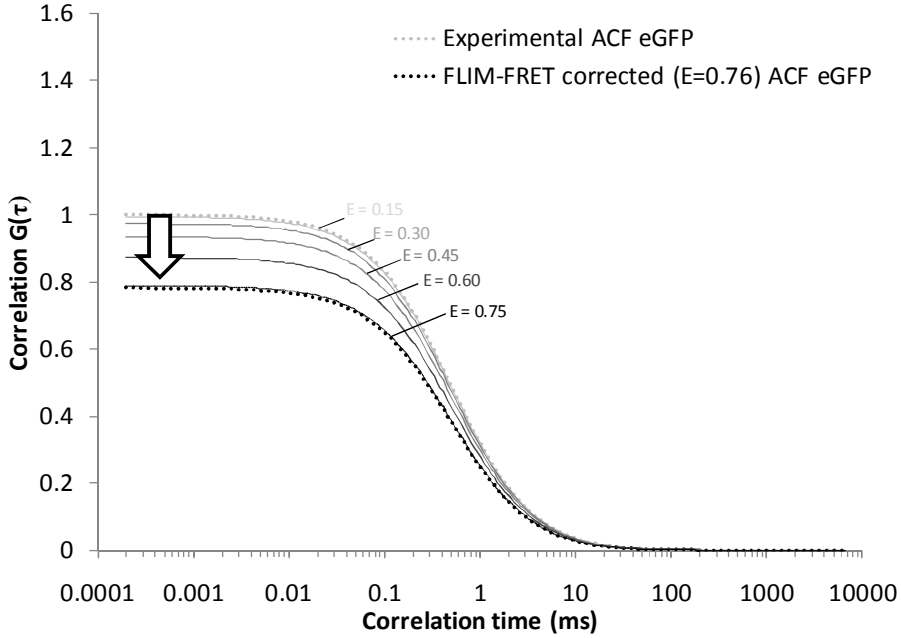

Figure 6.3 **The effect of the FRET correction on the eGFP ACF** –The simulated eGFP ACFs are calculated with Equation 6.2, with F_g and F_{gr} from the FLIM-FRET experiments. Arrow points from 'no correction' to 'experimentally determined correction'.

In summary, by quantifying the FRET between mCherry and eGFP in the fusion construct, we could correct the eGFP ACF to obtain an unbiased particle number. Since FRET lowers the brightness of the eGFP in a functional fusion protein, it lowers the observed particle number. The unbiased particle number is, due to the high FRET, 28% higher than the apparent particle number.

6.3.5. Correcting the mCherry autocorrelation function

From the corrected number of green particles, $N_{g,total,real}$, and from the fractions we can now calculate the number of 'only-green' particles, N_g, and the 'green-and-red' particles, N_{gr}. The red ACF has to be corrected for both cross-talk and FRET. If we assume that there are no 'only-red' particles, which is justified since eGFP did not show much photobleaching, it can be shown that the apparent red particle number is equal to:

$$N_{r,app} = \frac{\left(QN_g + Q(1-E)N_{gr} + (1+E)N_{r,total,real}\right)^2}{Q^2N_g + Q^2(1-E)^2N_{gr} + (1+E)^2N_{r,total,real}} \qquad \text{Equation 6.3}$$

Where N_g and N_{gr} are obtained from the green ACF and

$$Q = \frac{B_c}{B_r} = \frac{the\ red\ emission\ of\ the\ green\ fluorochrome}{the\ red\ emission\ of\ red\ fluorochrome} \qquad \text{Equation 6.4}$$

Since we assume no 'red-only' particles are present in our system, $N_{r,total,real}$ was approximated with N_{gr}, rescaled in the red confocal volume. Q was calculated by measuring the brightness of eGFP and mCherry in the red detection channel. First the ACF of eGFP was measured in the red detection channel upon single color 488-nm excitation and the apparent brightness was calculated (Equation 2.7). Subsequently, the ACF of mCherry was measured in the red detection channel upon single color 543-nm excitation and its apparent brightness was calculated. Q, the ratio of these apparent brightnesses, was then calculated to be ~0.2.

In Figure 6.4 we have simulated the effects of the cross-talk and FRET corrections on the ACF of mCherry for different values of Q and FRET. As can be seen in the top panel, cross-talk will lower the apparent mCherry ACF amplitude, since more particles are 'seen' in this channel when there is crosstalk. In absence of cross-talk, FRET would have no influence on the mCherry ACF, since only particles of brightness (1+E) are 'seen' in the red channel. However, as can be seen in the lower panel of Figure 6.4, in presence of cross-talk FRET will increase the apparent mCherry ACF amplitude. Importantly, under the experimental conditions of crosstalk (Q=0.2) and FRET (E=0.76), the calculated mCherry ACF coincides

perfectly with the experimentally measured mCherry ACF, justifying our assumption of no 'only-red' particles, i.e. $N_{r,total,real}=N_{gr}$, rescaled in the red volume.

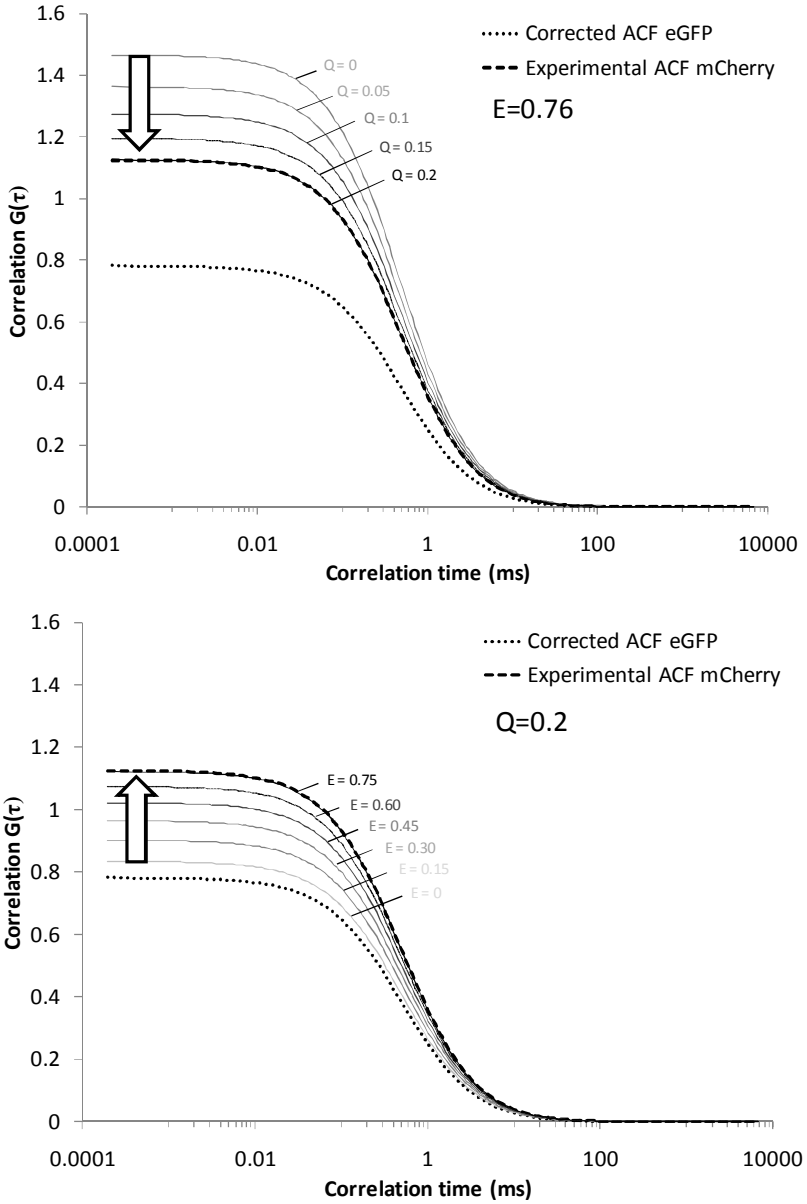

Figure 6.4 **The effects of cross-talk and FRET corrections on the mCherry ACF** – (top panel) The simulated mCherry ACFs are calculated with Equation 6.3 and with E=0.76, as experimentally determined. (bottom panel). The simulated mCherry ACFs are calculated with Equation 6.3 and with Q=0.2, as experimentally determined. The arrows point from 'no correction' to 'experimentally determined correction'.

In summary, by using the unbiased eGFP particle number and the fractions of green-only and green-and-red particles from the FLIM-FRET experiment, we could calculate an ACF for mCherry that was exactly the same as the measured ACF. Of concern, a 'red-only' population did not need to be included.

We thus have strong suggestions that the population indeed consist of a mixture of green-only and green-and-red particles, which can be explained through photobleaching part of the RFP population during the experiment. These simulations were also done for the mRFP-eGFP fusion, and gave comparable results.

6.3.6. Correcting the cross-correlation function

From the known 'green-only' and 'green-and-red' particle numbers, and with the proper crosstalk and FRET correction, we should be able to predict the amplitude of the cross-correlation function. We worked out Equation 2.32 to include both the contribution of FRET and crosstalk:

$$G_{CC}(\tau) = \frac{(1-E)^2 Q N_{gr} + (1-E)(1+E)N_{gr} + Q N_g}{\left(N_g + (1-E)N_{gr}\right)\left((1+E)N_{gr} + (1-E)Q N_{gr} + Q N_g\right)} \quad \text{Equation 6.5}$$

Again, no 'red-only' particles are assumed, so that this equation simplifies to $G_{CC}(\tau)=1/N_{g,total}$ in the absence of FRET. In Figure 6.5, we simulated the effects of FRET and crosstalk for different values of E and Q, including our experimentally determined values. As can be seen in the top panel of Figure 6.5, in the presence of FRET (E=0.76) but no crosstalk, the observed mCherry-eGFP CCF under the experimental conditions (F_g = 0.61 and F_{gr} = 0.39) would have an amplitude of only 0.27. Including cross-talk in the calculation of the CCF will increase this amplitude to 0.39, which is very close to the measured 0.44. Likewise, as illustrated in the bottom panel of Figure 6.5, in absence of FRET, the observed CCF is not biased by crosstalk (Equation 6.5) and its amplitude is equal to that of the green ACF. FRET clearly has a profound effect on the CCF amplitude and lowers it to a value that is very close to the experimentally observed value.

In summary, it was possible to predict the amplitude of the CCF of mCherry-eGFP with reasonable accuracy if we used our determined Q and E parameters and if the fractions of green-only and green-and-red particles are known. As expected, cross-talk increases the apparent CCF amplitude, and more importantly, FRET lowers the CCF amplitude considerably.

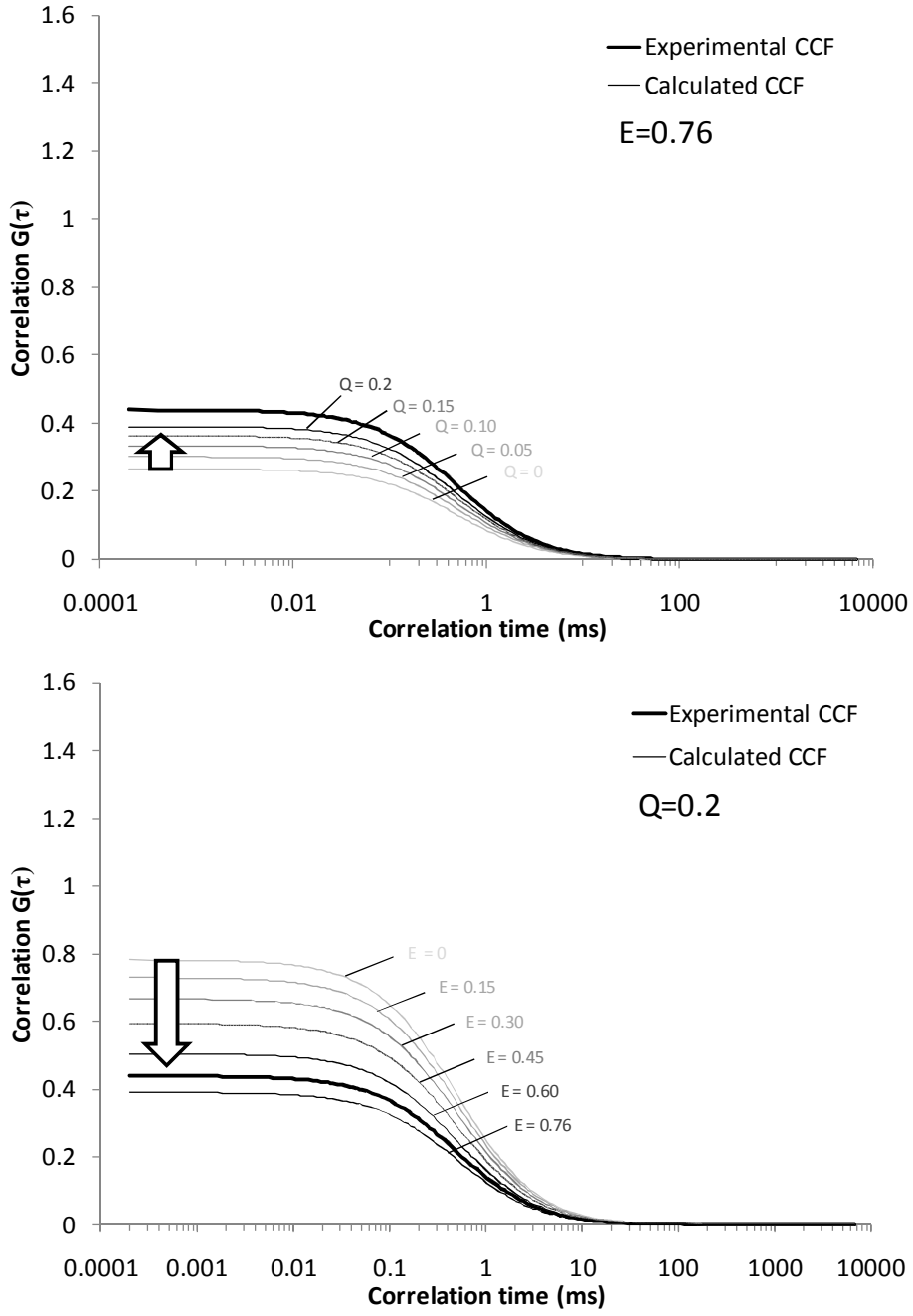

Figure 6.5 **The calculated and experimental CCF of mCherry-eGFP** – The simulated CCFs are calculated with Equation 6.5 and with either (top panel) E=0.76 or (bottom panel) Q=0.2. The arrows point from 'no correction' to 'experimentally determined correction'.

6.4. Discussion

In this chapter we wanted to verify if FCCS can be used to study the co-diffusion of a green (eGFP) and a red (mRFP1 or mCherry) fluorescent protein inside living cells in a very simple model system consisting of a genetic fusion of the two proteins. The principle for measuring and quantifing protein-protein interactions with FCCS on the Zeiss ConfoCor2 microscope is based on the independent excitation and detection of the two fluorochromes that are used

to label the proteins. Non-idealities of our experimental setup and principle do exist and can be divided into fluorochrome and instrument related non-idealities.

6.4.1. Fluorochrome non-idealities

With respect to the fluorochromes, we chose the superior fluorescent protein couples that were available at the time, eGFP and either one of two monomeric RFPs (mRFP1 or mCherry). It could be argued that eGFP and mRFP are, from a spectral point of view, not the right choice for performing two-color FCCS, because their small Förster distance (~5 nm) renders them good FRET partners. However, the emission spectra of cyan or even blue fluorescent proteins still show considerable overlap with the emission spectrum of common mRFPs. Furthermore, no far-red fluorescent proteins exist that are compatible with FCCS. Enhanced green fluorescent protein, already engineered in 1995 (Heim et al., 1995), to date still has excellent properties for quantitative measurements of protein dynamics; high brightness and photostability and a monomeric stoichiometry. However, it must be stated that eGFP has been shown *in vitro* to be able to enter short-lived transient dark states such as the triplet process and pH-dependent and –independent protonation (Haupts et al., 1998; Widengren et al., 1999a), rendering the protein non-fluorescent for periods up to 300 μs, which is in the time range of the diffusion dynamics and could potentially compromise quantitative measurements. In 2004, engineering of monomeric red fluorescent proteins with a good brightness opened the way for quantitative two-color measurements of dynamics. However, some properties of RFPs limit their use in quantitative spectroscopy. First, up to date, all red fluorescent proteins have been shown to enter microsecond-to-second time scale light enhanced dark states (Heikal et al., 2000; Malvezzi-Campeggi et al., 2001; Schenk et al., 2004b; Hendrix et al., 2008). In 0, the photophysics of mRFPs will be investigated in more detail. Secondly, slow chromophore maturation has been reported in some mRFPs. DsRed for example, the parent protein from which mRFP1 and mCherry were constructed, has a maturation time of almost a day (Shaner et al., 2004). Although mCherry has been shown to have quite a fast (20 min.) maturation time, we do not rule out this possibility. Thirdly, it could be argued that an error in transcription or translation for the fusion constructs exists. However, when the position of eGFP and the RFP are switched, or when a modified amino acid linker was inserted between the two proteins, similar results were obtained. Finally, mCherry and even more its predecessor mRFP1, are quite sensitive to photobleaching. Photobleaching irreversibly removes red molecules and partially explains the lower average

particle number of red molecules. Improvements in RFP fluorophore photostability have recently been achieved with a new RFP, TagRFP-T, that has an estimated 40-fold improved photostability and 3-fold higher brightness compared to mRFP1 (Shaner et al., 2008).

6.4.2. Instrumental non-idealities

With respect to the instrumental non-idealities, two-color excitation was used for these experiments. We are confident that the confocal volumes are concentric in the XY direction since the two excitation beams, 488 nm and 543 nm originate from the same optical fiber in the fiber coupling unit (FCU) of the Zeiss ConfoCor2 microscope, and follow the same optical path. On the other hand, the wavelength dependence of the light refraction in the objective could cause the confocal volumes created by the 488-nm and 543-nm lasers to be non-concentric in the Z-direction, resulting in a smaller cross-correlation volume than expected (Figure 6.6). Since the CCF amplitude is directly proportional to the cross-correlation volume (Equation 6.5), a lower CCF amplitude is obtained than expected. Although incomplete volume overlap does put a limit on the maximum CCF amplitude possible, in our case the effect was rather small and correcting for this in the analysis of FCCS is possible when performing a detailed characterization of the microscopic setup (see section 3.3.6). Furthermore, the observation of low CCF amplitude for a fusion protein has been observed even for single-wavelength (514 nm) excitation FCCS on fluorescent proteins (Liu et al., 2007).

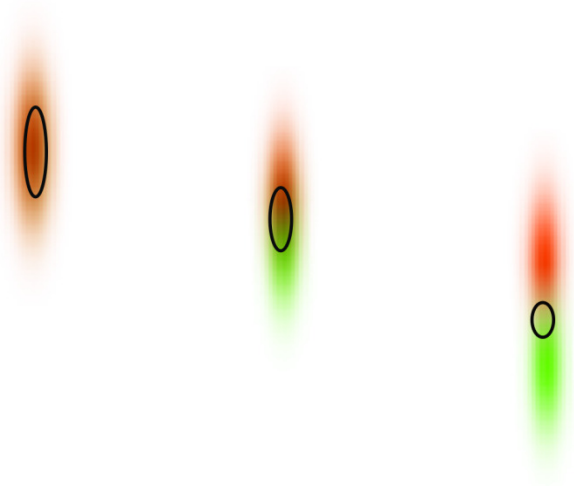

Figure 6.6 **Axial chromatic aberration effect on FCCS** – From left to right: the less the two excitation spots, shown in red and green, overlap in Z-direction, the smaller the effective cross-correlation volume is, shown by the gray ellipsoid.

Secondly, since the microscope is equipped with 'continuous wave' (CW) lasers, we simultaneously excited the green and red fluorescent proteins. As seen earlier (sections 2.3.4.2 and 2.3.4.3), this measurement principle has two main consequences when using fluorescent proteins: cross-talk of the green fluorochrome into the red channel and the possibility of Förster resonance energy transfer between the fluorochromes. In the experiments performed here all the parameters (crosstalk parameter Q, protein fractions F_g and F_{gr} and FRET efficiency E) needed to correct for these non-idealities could be extracted from control experiments. It has been shown that cross-talk can be eliminated by using temporal separation of the excitation, either by using acousto-optical tunable filters or modulators, or by using picoseconds pulsed lasers instead of CW lasers. Furthermore, since green and red fluorochromes are not excited simultaneously with these techniques, the red ACF and CCF are not influenced by FRET in the red detection channel. This would simplify the analysis of FCCS considerably and would likely make the technique more robust for quantitative analysis of protein-protein interactions.

6.5. Conclusion

Fluorescent protein cross-correlation spectroscopy has the potential of becoming the biophysical tool of choice for quantifying protein-protein interactions in cells, because the absolute concentrations of both partners and their complex can directly be measured. We have shown for a model system that FCCS measurements can be analysed quantitatively in the presence of dark states, cross-talk and FRET and have discussed the different fluorochrome and experimental setup related issues that can be optimized to render the technique more robust. Especially FRET proved to profoundly influence the correlation functions and should better be avoided. Future work will consist in searching for a better methodology for FCCS, both by applying pulsed-interleaved excitation or temporal laser switching FCCS, to avoid crosstalk and acceptor FRET artefacts, and by using new fluorescent proteins with a larger spectral separation, such as BFP or CFP in combination with RFP.

Chapter 7. Dark states in monomeric red fluorescent proteins

A fluctuating fluorescence signal usually means molecules are diffusing into and out of the confocal volume… or doesn't it? We have already seen that fluctuations due to noise, although not correlated in time, have a negative effect on FCS, and that fluctuations due to the triplet conversion, that do correlate in time, contribute to the autocorrelation function in the fast time scale. If one wants to study diffusion and concentrations, then a detailed understanding of the autcorrelation behavior of fluorochromes is necessary. In this chapter we will focus specifically on a certain category of fluorochromes, the monomeric red fluorescent proteins. Being an ideal partner of green fluorescent proteins for FCCS measurements, we have to characterize in detail how these proteins behave in a fluorescence fluctuation experiment. We discuss experiments performed in a controlled *in vitro* environment and provide detailed and useful information about the photophysics of these proteins.

Part of this chapter has been published in the following article:
Hendrix,J., Flors,C., Dedecker,P., Hofkens,J., and Engelborghs,Y. (2008). Dark states in monomeric red fluorescent proteins studied by fluorescence correlation and single molecule spectroscopy. Biophys. J. 94, 4103-4113.

7.1. Introduction

The green fluorescent protein was discovered in 1961 in the hydrozoan *Aequorea victoria* (Shimomura, 2005). Fluorescence microscopy in live cells has been revolutionized since the advent of GFP as a genetically encoded marker in 1992 (Prasher et al., 1992). In the following years, variants of the GFP with different absorption/emission maxima and higher brightness

and stability have created a fluorescence palette ranging from blue (eBFP) to yellow (eYFP, Citrine). In 1999 this palette was extended to the red by the discovery of red fluorescent proteins (RFPs) in anthozoans like *Discosoma* sp. (DsRed) and in 2002 also in *Entacmaea quadricolor* (eqFP611) (Matz et al., 1999; Wiedenmann et al., 2002). Irrespective of their color, all fluorescent proteins have a β-barrel tertiary structure, protruded by an internal α-helix. After folding of the protein, the light absorbing and emitting fluorophore spontaneously forms by chemical reactions of amino acids in this internal α-helix, as illustrated for DsRed in Figure 7.1. Importantly, the physicochemical properties of the reacting amino acids, together with the surrounding amino acids forming the fluorophore pocket, are the basis for both the color and photophysical properties of the fluorescent protein.

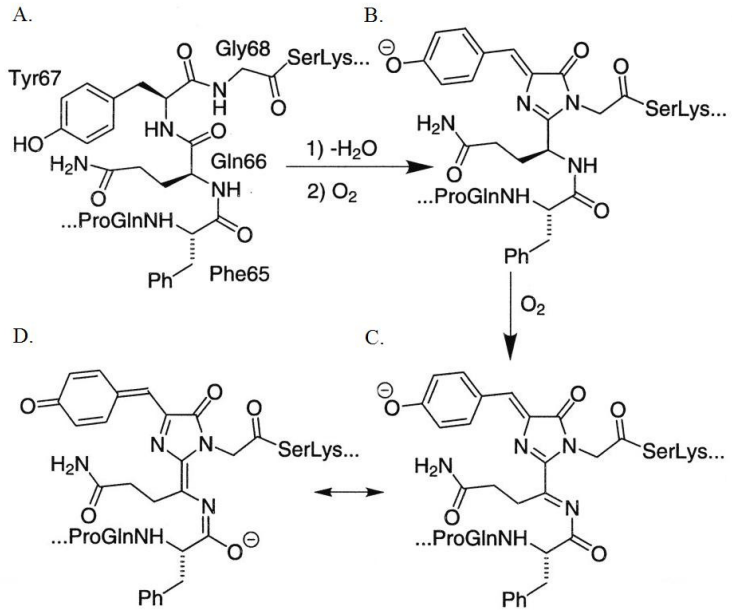

Figure 7.1 **Formation of the DsRed fluorophore** – (A) the DsRed polypeptide in the right orientation. (B) the immature green fluorophore. (C) The mature red fluorophore. (D) Resonant form of (C) (Campbell et al., 2002).

Being an obligate tetramer, DsRed was engineered in 2002 by Campbell et al. into mRFP1, the first true monomeric red fluorescent protein (Campbell et al., 2002). mRFP1 contains a total of 33 amino-acid mutations, but is structurally still very stable (Stepanenko et al., 2004). The chromophore in mRFP1 matures much faster than in DsRed, is less bright and its photostability is lower than DsRed, but still comparable to eGFP (Campbell et al., 2002).

In 2004, further engineering of mRFP1 led to the mFruits family that covers the yellow (mHoneydew) to dark-red part (mPlum) of the visible spectrum (Shaner et al., 2004; Wang et al., 2004). Of the mFruits, mCherry is considered to be the superior monomeric RFP (Shaner

et al., 2005). It has a good photostability and even faster maturation but still is only moderately brighter than mRFP1. A brighter version of mRFP1 exists (Q66T), but at the cost of a 34-nm blue shift of the emission spectrum (Jach et al., 2006). Recently, a new mRFP (TagRFP) has been created from eqFP578, but the emission spectrum still is 24-nm blue-shifted with respect to mRFP1 (Merzlyak et al., 2007). mRFP1 has also been converted into a photoactivatable protein (Verkhusha and Sorkin, 2005b).

mRFPs are in great demand for experiments in live cells because of their interesting properties. First, their red emission leads to less scattering and less background fluorescence detection. Second, their emission spectrum can be well separated from eGFP, which makes them interesting for dual colour applications. Moreover, when fused to a certain protein, their monomeric structure preserves the natural function and allows for FRET measurements. However, previous photophysical studies of some RFPs have shown the presence of light-induced dark states that can complicate the interpretation of the results of the cellular measurements (Heikal et al., 2000; Malvezzi-Campeggi et al., 2001; Schenk et al., 2004a).

In 2000, Heikal et al. observed with FCS that DsRed in solution shows a large contribution of an excitation intensity-dependent (but not pH-dependent) flickering in the sub-ms timescale (Heikal et al., 2000). Subsequently, Malvezzi-Campeggi et al. proved that this flickering represents the transition between three interconvertible states (a red, a far-red and a dark state) (Malvezzi-Campeggi et al., 2001). Schenk et al. confirmed the observations for DsRed and observed a similar process in a different RFP, eqFP611 (Schenk et al., 2004a). The flickering was suggested to be a consequence of conformational rearrangements around the chromophore, such as photoisomerization and/or changes in the hydrogen-bonding network. Both for DsRed and eqFP611, Raman spectroscopy later showed that photoisomerization is indeed involved in the photodynamics of these RFPs (Habuchi et al., 2005; Loos et al., 2006). In some green fluorescent protein mutants such as eYFP and Citrine (Heikal et al., 2000; Schwille et al., 2000), a pH- and intensity-dependent flickering was also found and assigned to fast protonation-deprotonation.

In this chapter, we focus on the monomeric red fluorescent proteins mRFP1, mStrawberry and mCherry and on the presence of light-induced flickering. We have characterized this process with fluorescence correlation and single-molecule spectroscopy, and we have explored the effects of viscosity, excitation intensity and pH on this process. We have combined our experimental results with information based on known X-ray structures of

some RFPs, and we provide insight into the consequences that the chromophore environment has on the observed fluorescence properties. Also, we draw the attention to the implications of the fast flickering of RFPs when performing quantitative FCS measurements in live cells.

7.2. Materials and Methods

7.2.1. Proteins and buffers

The pRSET plasmids coding for His-tagged mRFP1, mCherry and mStrawberry were a kind gift of Dr. Roger Y. Tsien (HHMI-UCSD, La Jolla, CA). After transformation of the plasmid in *E. coli* BL21 cells, the cells were grown at 37 °C to an optical density of 0.6 after which protein overexpression was induced during 3 h with 1 mM IPTG. After sonication of the culture, the proteins were purified using gravity flow Ni^{2+}-affinity chromatography (Protino Ni-TED, Macherey-Nagel GmbH & Co. KG, Düren, Germany).

The protein purity was checked with SDS-PAGE. On this gel, three bands could be observed, corresponding to native and cleaved protein. The cleavage occurs in the chromophore and is caused by a partial hydrolysis of the main chain acylimine linkage ((Gross et al., 2000; Campbell et al., 2002; Verkhusha and Sorkin, 2005b) and data not shown). The extent of the cleavage does not increase over time, suggesting that the cleavage occurs only in a subpopulation which may represent poorly folded protein. This suggestion is underpinned by the fact that in the photoactivatible PA-mRFP1-1 a larger extent of cleavage occurs than in mRFP1 (Verkhusha and Sorkin, 2005a), which might be due to the mutations in the chromophore environment. The cleavage irreversibly changes the spectral properties of the RFP, converting it from a red to a green-like protein. The residual fluorescence of the green-like protein, if any, will not influence our measurements, since we are using a red emission filter. Most of the experiments were performed in phosphate buffered saline (PBS, pH 7.4). The pH buffer consisted of 50 mM of phosphate, citrate and glycine adjusted from pH 3-12 (with 0.5 increments) with 3 N of NaOH. If a higher viscosity was needed, glycerol was added to the buffer, and the pH was adjusted with NaOH containing the same concentration of glycerol. Samples for single-molecule measurements were prepared by spin-coating the proteins (~10^{-11} M) in PBS containing 1% (wt) polyvinyl alcohol (PVA) on a clean cover glass at 3000 rpm.

7.2.2. Absorption and emission spectra

Absorption spectra were measured with a Shimadzu UV-1601PC spectrophotometer (Shimadzu GmbH, Duisburg, Germany). Excitation and emission spectra were taken with a PTI QuantaMaster fluorometer (Photon Technologies International, West Sussex, UK). For the excitation spectra, the excitation monochromator bandwidth was 2 nm and the emission monochromator to 650±10 nm. For the emission spectra, the excitation monochromator was set to 540±10 nm and the emission monochromator bandwidth was set to 2 nm.

7.2.3. FCS in solution

For FCS measurements in solution the concentrated protein was diluted to 1 nM in the appropriate buffer. The measurements were performed on a commercial ConfoCor2 system (Carl Zeiss, Jena, Germany).

A 543-nm HeNe laser was used to excite mRFP1, with the acousto-optical tunable filter set to 3-100%, corresponding to an intensity of 2-122 kW/cm^2 in the confocal spot (as measured with a light power meter). The excitation light was reflected by a HFT543 dichroic mirror and focused in the sample through a C-Apochromat 40×/1.2NA water immersion objective. The fluorescence light was filtered by a LP560 longpass filter and was detected on an avalanche photodiode through a 78 μm pinhole. For the different experimental conditions (intensity, pH,...) 10 measurements of 20 seconds were performed and the average autocorrelation curve was calculated. The measurements were analyzed in Igor Pro 5 (Wavemetrics, Inc., Oregon, USA) by means of global analysis with Equation 2.17, including 1 or 2 dark states.

7.2.4. Single-molecule fluorescence spectroscopy

For single-molecule experiments, excitation at 543 nm (8 MHz, 1.2 ps FWHM) from the frequency doubled output of an optical parametric oscillator (GWU Lasertechnik, Erftstadt-Friesheim, Germany) pumped by a Ti:Sapphire laser (Tsunami, Spectra Physics) was directed into an inverted microscope (Olympus IX 70, Tokyo, Japan) and focused onto the sample through an oil immersion objective (60×/1.4NA, Olympus). Fluorescence was collected through the same objective and sent to an avalanche photodiode (SPCM-AQR-15, PerkinElmer, Whaltham, MA). For the single molecule experiments, the excitation intensity was 120 nW, which corresponds to ~ 261 W/cm^2. Time-resolved data were collected with a time-correlated single photon counting (TCSPC) card (SPC 630, Becker & Hickl, Berlin,

Germany) operated in first-in-first-out mode. The detailed description of the setup and the data acquisition process have been published previously (Vosch et al., 2003).

7.2.5. Ensemble time-resolved fluorescence

Ensemble time-resolved fluorescence measurements were performed using the TCSPC technique (Maus et al., 2001), with excitation at 543 nm using the same source as above, in a cuvette with a path length of 1 cm and an optical density of 0.1 at the absorption maximum. Fluorescence was detected under magic angle geometry, by means of a cooled microchannel plate photomultiplier (Hamamatsu R3809U). Fluorescence histograms of the sample and of the instrument response function were collected in 4096 channels, until they typically reached 10^4 counts in the peak channel. The total width at half-maximum of the instrument response function was ~ 40 ps. Single-molecule and ensemble time-resolved fluorescence spectroscopy measurements were performed by Dr. Cristina Flors at the Laboratory for Photochemistry and Spectroscopy, under the supervision of Prof. Johan Hofkens.

7.3. Results

7.3.1. Fluorescence correlation spectroscopy of mRFP1 and mFruits

We applied FCS to discover if dark-state formation also occurs in mRFP1, mCherry and mStrawberry, as found for the other RFPs described in the introduction. Since the viscosity of the medium only affects the translational diffusion, we can exploit it for the identification of non-diffusion-related processes. Whereas at 24 kW/cm² in aqueous buffer there are no distinct components visible in the autocorrelation curve (Figure 7.2, left curve), in 50% (v/v) glycerol two different components are clearly visible (Figure 7.2, right curve). The fastest component can be adequately described by an exponential blinking term in the fitting function and the slowest is best described by a translational diffusion term. The relaxation time of the fast component is much faster than the expected time necessary for a typical monomeric GFP to diffuse through the confocal volume, confirming its identity as a non-diffusion related process. The fraction of flickering for mRFP1 and for mCherry molecules at 24 kW/cm² is similar, 46±1% and 43±2% respectively. At the same I_{exc} mStrawberry showed a markedly increased fraction, 56±1%. Furthermore, the flickering also experienced a small viscosity effect. It should be noted, however, that glycerol changes the refractive index of the solution, which also has an effect on the autocorrelation curve (Enderlein et al., 2005).

Nevertheless, this experiment proves that also in these mRFPs dark-state formation occurs in a timescale of tens of μs.

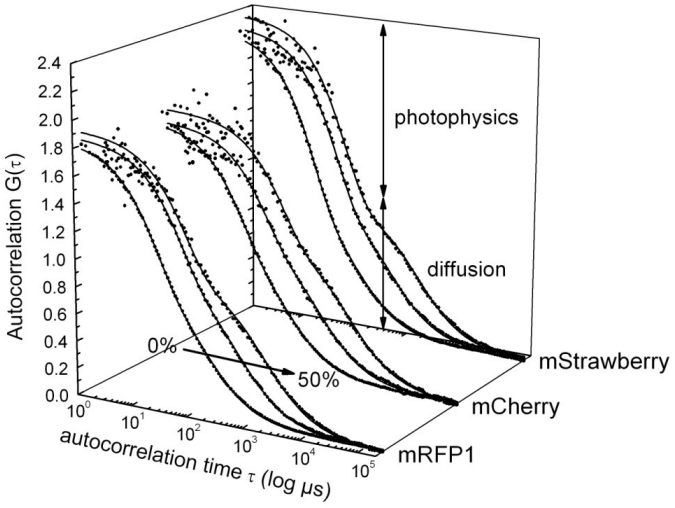

Figure 7.2 **Experimental FCS curves of the mRFPs at I_{exc} = 24 kW/cm² normalized to the diffusion component** – In each group of 3 curves the concentration of glycerol was 0, 25, 50% (v/v). From left to right: mRFP1, mCherry, mStrawberry. The solid lines are global fits with Equation 2.17, where the fraction of the flickering was linked within one set.

7.3.1.1. Intensity dependence

To determine if dark-state formation is light-induced, we performed FCS measurements at different excitation intensities, ranging from 2 to 122 kW/cm² (Figure 7.3 *A*). First, we determined the diffusion coefficient of mRFP1 in aqueous solution. The brightness of mRFP1 was not linearly dependent on I_{exc} at high excitation intensity (data not shown), which can be attributed to optical saturation of the fluorescent protein (Enderlein et al., 2005). This optical saturation in combination with a larger contribution of flickering processes at higher I_{exc} caused the apparent D to vary with I_{exc}. By measuring at different I_{exc} and extrapolating to I_{exc} = 0 (Figure 7.3 *B*), we obtained a value of D = ~56 μm²/s, which is in good agreement with the value found for eGFP. Two-focus FCS might provide a confirmation for the value we obtained, since in this technique D is measured relative from one focus to another, and is less prone to optical artifacts (Dertinger et al., 2007b).

Figure 7.3 *C* and *D* show respectively the relaxation time and the fraction of the non-diffusion-related process(es). In the low I_{exc}-regime, the fraction and relaxation time of the flickering are clearly light-dependent (squares in Figure 7.3 *C* and *D*), meaning that the path

towards the dark state is favored, i.e., the dark state formation is indeed light-induced. In the higher I_{exc}-regime the fraction of the flickering appears to become constant.

In addition, a second exponential term had to be included in the fitting. As an illustration, a fit of the FCS curve at the highest I_{exc} with one exponential function is also shown in Figure 7.3 *A*. The relaxation time of the second exponential term could be kept constant and the fraction showed a linear increase (circles in Figure 7.3 *C* and *D*), even in the region where the other process appeared to have become saturated. This second process might represent another light-induced state, e.g. triplet conversion (Widengren et al., 1999b; Jung et al., 2005). Triplet lifetimes of red fluorescent proteins such as HcRed have been estimated at a few microseconds (Cotlet et al., 2006), similarly to our observations. Due to the low triplet quantum yield in GFPs (Jimenez-Banzo et al., 2008) the exponential term that accounts for its formation only becomes apparent at higher I_{exc}, as has already been observed for other GFPs (Haupts et al., 1998; Widengren et al., 1999b). These measurements thus suggest that there are (at least) two different dark-states for mRFP1 When mCherry and mStrawberry were measured, a similar trend was observed (Figure 7.4 and Figure 7.5).

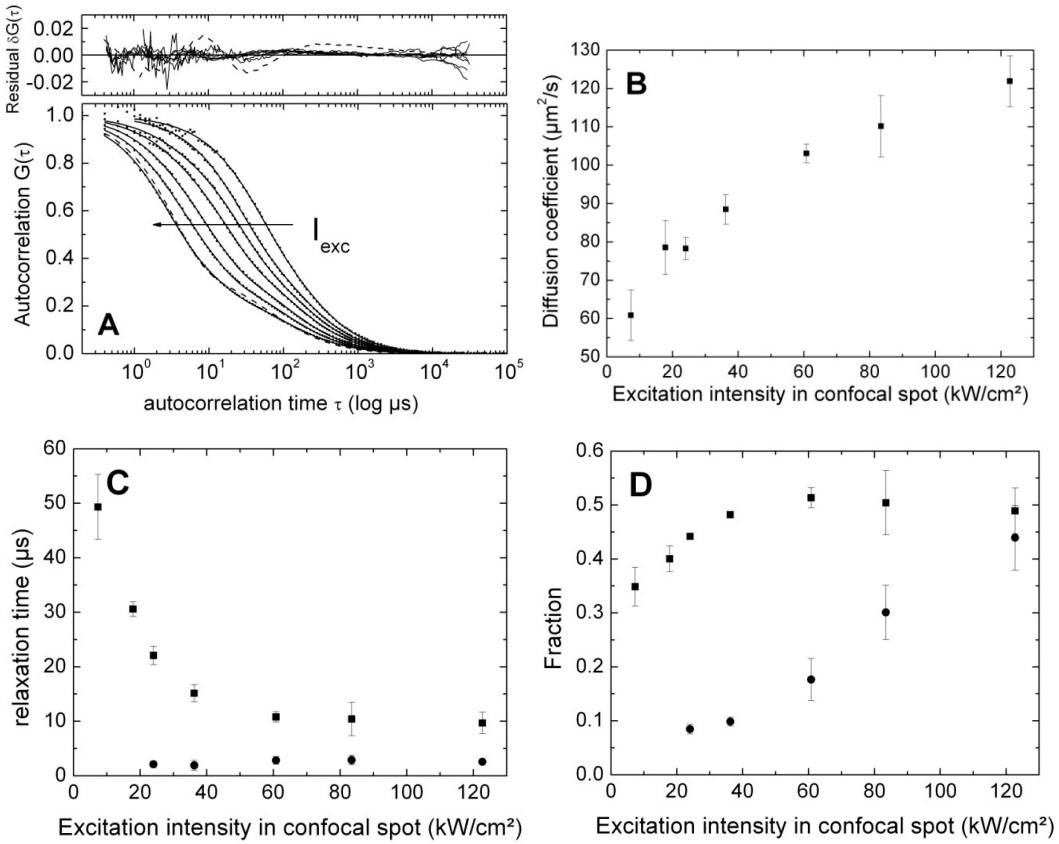

Figure 7.3 **Intensity dependence of the ACF of mRFP1** (A) Experimental autocorrelation curves of mRFP1 in aqueous buffer at excitation intensities ranging from (right to left) 2-122 kW/cm². The curves were normalized to $G(0) = [N.(1-F_1).(1-F_2)]^{-1}$ to emphasize the effect of I_{exc} on the fast component of the curve. It is worth to note that normalization in this manner causes the $G_D(0)$ (Equation 2.17) to decrease at increasing I_{exc}, even though the concentration stays constant. For the highest I_{exc}, both a one- (continuous line) and two- (dashed line) exponential fit are shown. (B) Apparent diffusion coefficient after fitting with Equation 2.17. (C) Relaxation times of the flickering (squares) and triplet (circles) processes. (D) Fraction of the fast processes when G(0) is normalized to 1.

Moreover, for all three mRFPs we find that for the slower flickering process the associated dark fraction does not tend to zero in the limit of zero excitation intensity (Figure 7.3 D, Figure 7.4 D and Figure 7.5 D).

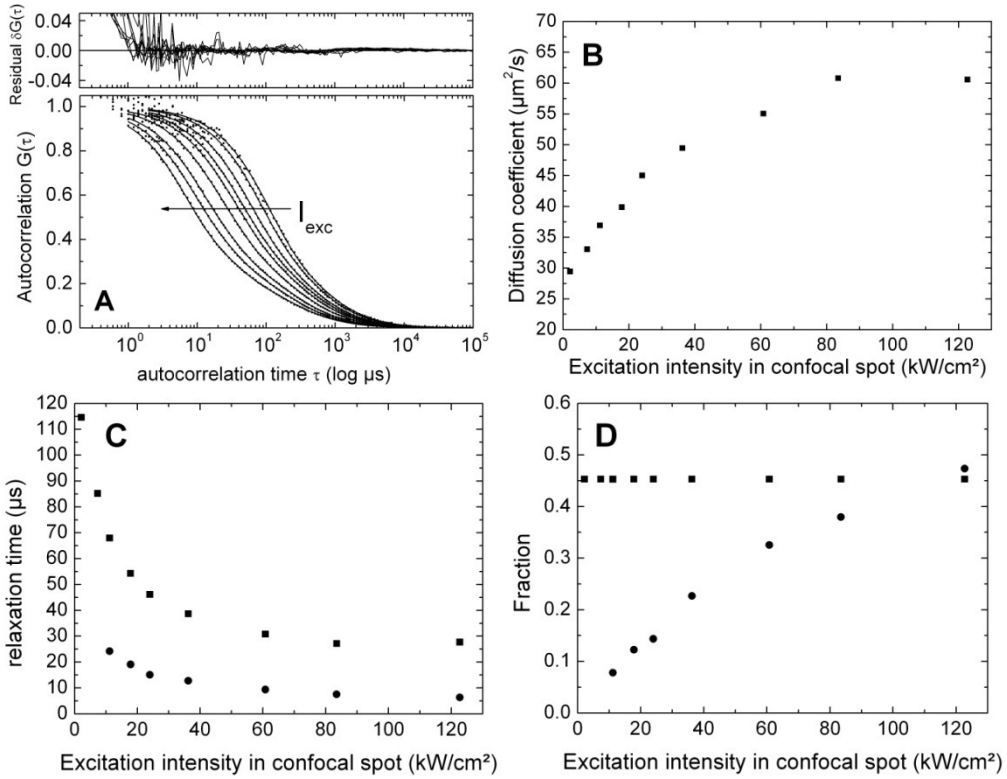

Figure 7.4 **Intensity dependence of the ACF of mCherry** (A) Experimental autocorrelation curves of mCherry in aqueous buffer at excitation intensities ranging from (right to left) 2-122 kW/cm². The curves were normalized to $G(0) = [N.(1-F_1).(1-F_2)]^{-1}$ to emphasize the effect of I_{exc} on the fast component of the curve. The measurements on mCherry did not show a clear trend in F_1. When the latter parameter was kept constant, this resulted in a intensity dependence of both relaxation times. Due to the relatively small difference in both relaxation times and diffusion time, especially at high I_{exc}, it is difficult to say which fitting method yields the more realistic parameters. (B) Apparent diffusion coefficient after fitting with Equation 2.17. (C) Relaxation times of the flickering (squares) and triplet (circles) processes. (D) Fraction of the fast processes when G(0) is normalized to 1. Because there was no clear trend in F_1, it was globally linked.

Therefore we conclude that one of the two dark states observed with ensemble FCS is populated even in the absence of excitation light, but that its formation is accelerated through irradiation, while the other dark state is only occupied when excitation light is present.

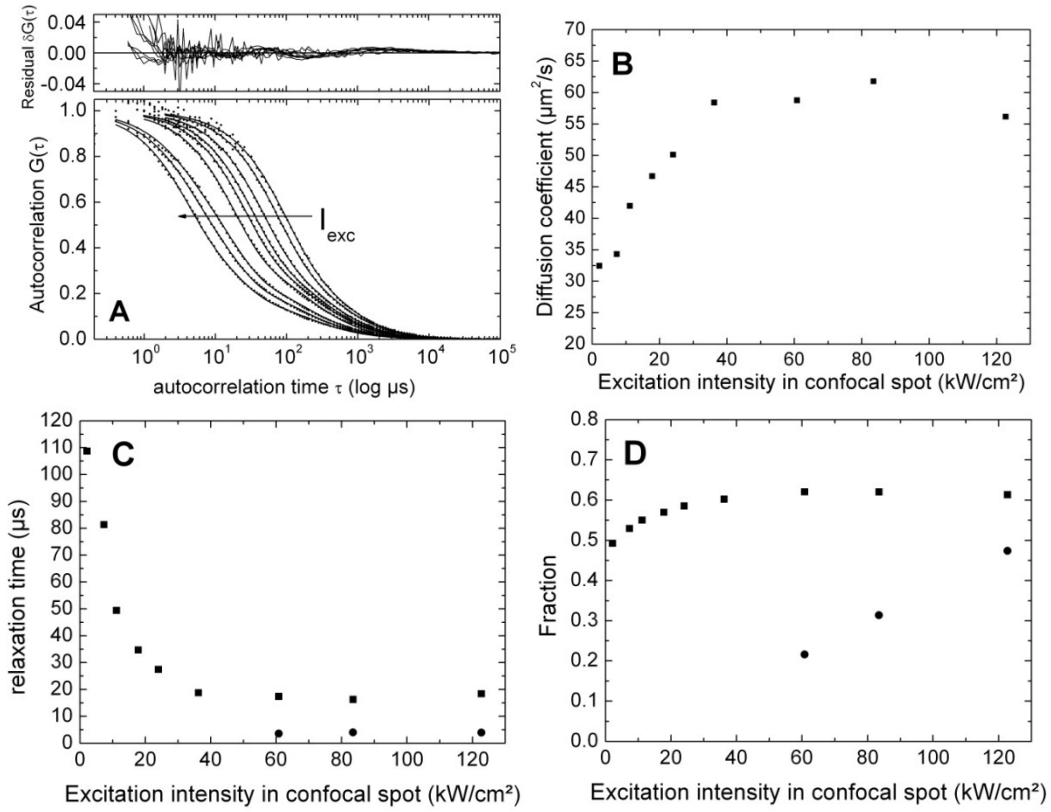

Figure 7.5 **Intensity dependence of the ACF of mStrawberry** (A) Experimental autocorrelation curves of mStrawberry in aqueous buffer at excitation intensities ranging from (right to left) 2-122 kW/cm². The curves were normalized to $G(0) = [N.(1-F_1).(1-F_2)]^{-1}$ to emphasize the effect of I_{exc} on the fast component of the curve. (B) Apparent diffusion coefficient after fitting with Equation 2.17. (C) Relaxation times of the flickering (squares) and triplet (circles) processes. (D) Fraction of the fast processes when G(0) is normalized to 1.

7.3.1.2. pH-dependence

We explored the effect of varying pH on the dark-state formation in mRFP1 and mFruits, since it was shown for other fluorescent proteins that the pH can play a role in this process. FCS measurements of mRFP1 were performed at different pH values in 50% (v/v) glycerol (see Materials and Methods), to allow for an easy visual inspection of the autocorrelation curves (Figure 7.6 *A*).

To emphasize the effect of pH on the flickering dynamics, the autocorrelation curves were normalized with respect to the translational diffusion part. The fraction of mRFP1 molecules that is, on average, in a dark state shifted from 50±1% to 64±1% when increasing pH from 7 to 12 and the associated relaxation time shifted from 98.6±3.0 µs to 112.3±4.2 µs, indicating that mRFP1 is more often and longer in a dark state (since $1/\tau_{relaxation} = k_{relaxation} = k_{bright \rightarrow dark} + k_{dark \rightarrow bright}$). In mCherry and Strawberry, a similar increase in the fraction and relaxation

times were observed (Figure 7.6 *C*). In an acidic environment (pH ≤ 5) a different fast process appears in the autocorrelation curve (data not shown).

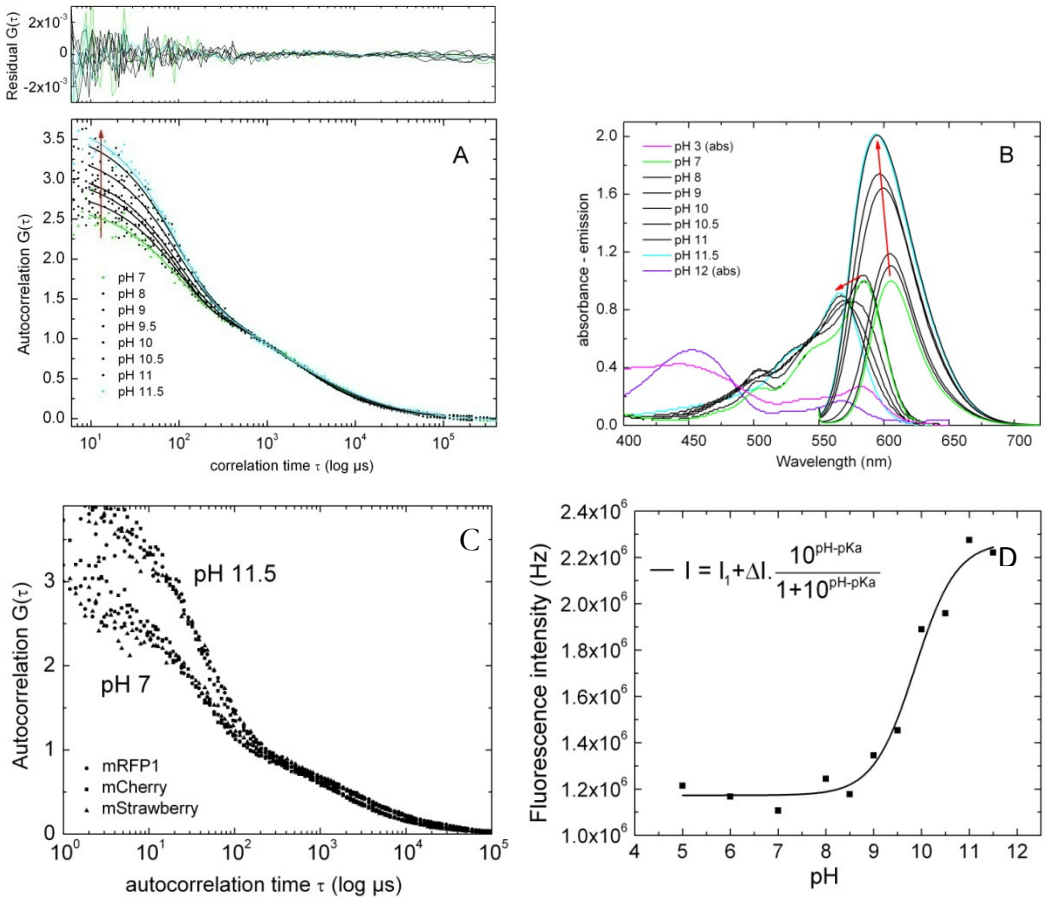

Figure 7.6 **pH dependence of the ACF of mRFP1** (A) Experimental autocorrelation curves of mRFP1 at I_{exc} = 61 kW/cm² in pH buffer supplemented with 50% (v/v) glycerol. The autocorrelation curves were normalized to the diffusion part of the curve. Normalization in this manner emphasizes the effect of pH on the flickering. The red arrow indicates the increase of the fraction of the flickering upon increasing pH from 7 to 11.5. (B) Absorption and emission spectra of mRFP1 at varying pH. The red arrow indicates the hypsochromic shift of the spectra and the increase of the quantum yield of mRFP1. (C) Comparison of the pH-dependence of the mRFP1- (squares), mCherry- (circles) and mStrawberry-autocorrelation curves (triangles) at I_{exc} = 61 kW/cm² in pH buffer supplemented with 50% (v/v) glycerol. The autocorrelation curves were normalized to the diffusion part of the curve. Normalization in this manner emphasizes the effect of pH on the flickering. (D) pH-titration of the emission intensity of mRFP1 at λ_{max}. I = fluorescence intensity of the sample, I_1 = fluorescence intensity when Glu-215 is 100% protonated and $\Delta I = I_2 - I_1$ the difference in fluorescence intensity when Glu-215 is deprotonated or protonated.

This process has already been described for eGFP and likely represents the protonation equilibrium of the hydroxyl moiety of the modified Tyr66 residue in the chromophore (Haupts et al., 1998). Since the chromophore is only fluorescent in the anionic state, protonation of Tyr66 at low pH quenches the fluorescence, giving rise to a dark state in the autocorrelation curve. Interestingly, although DsRed shares the same chromophore as mRFP1, the dark states of DsRed are not pH-dependent (Heikal et al., 2000). To investigate if the effects observed in FCS are related to changes in the steady-state absorption and emission properties, the pH-dependence of mRFP1 absorption, emission and excitation spectra were measured from pH 3 to 12 (Figure 7.6 *B*). The absorption maximum of mRFP1 shifted from 584 to 567 nm when changing pH from 7 to 11, which is in good agreement with measurements on mCherry (Shu et al., 2006). Below pH 7 the absorption spectrum did not change significantly. In the emission spectra, a similar blue shift from 605 to 595 nm was observed when increasing from pH 7 to 11 (Figure 7.6 *B*). More interestingly, when raising the pH from 7 to 11 the intensity of fluorescence drastically increased ($I_{\lambda,max}$ increased by a factor ~ 2.4). The shift of the emission intensity with pH fitted well with a model for a monoprotic acid/base sytem (Figure 7.6 *D*) and a pKa of 9.9 ± 0.1 was obtained for the process, which is in close accordance with values reported on mCherry and mStrawberry. The apparent fluorescence quantum yield of the chromophore thus rises as a result of a different protonation state of the chromophore environment. This brightness increase cannot be due to a decrease in the fraction or relaxation time of the dark state, since our FCS data show that at high pH the dark state lives longer and occurs more frequently. On the other hand, these observations can be explained based on the crystallographic information about mRFPs at different pH values.

E.g. for mStrawberry, a significant drop of the twist and tilt angles of the chromophore was observed when the pH was increased from 9.5 to 10.5 (Shu et al., 2006), in very good agreement with our observations on mRFP1. Chromophore co-planarity is a very important factor in the fluorescence efficiency of GFPs (as is discussed further on in this paper). Similar effects were seen in the excitation spectrum of mRFP1 (data not shown).

At pH 3 and 12 all fluorescence disappeared, indicating the denaturation of the protein. At these extreme values of pH the chromophore still showed a wide absorption peak in the blue (Figure 7.6 *B*), as has been reported for other fluorescent proteins (Gross et al., 2000).

7.3.2. Single-molecule spectroscopy of immobilized mRFP1

We used single-molecule spectroscopy of immobilized mRFP1 in order to study the fast on-off dynamics without the contribution of diffusion. A typical single-molecule intensity trace of mRFP1 in a PVA matrix is shown in Figure 7.7 *A* and *B*. Frequent on-off blinking is clearly observed in several timescales, even up to seconds, revealing a reversible transition from a bright to a dark state, similar to our observations in the FCS experiments. The flickering was found for the large majority of the molecules measured (>90%). In general, the emission intensity level is not stable due to this blinking, but there are no evident sequential photobleaching steps, which is consistent with the monomeric form of the protein. For comparison, immobilized DsRed showed several intensity levels as expected from its oligomeric form. In addition, on-off blinking was also present as reported before (Garcia-Parajo et al., 2001; Cotlet et al., 2001a; Cotlet et al., 2001b), especially at the lower intensity levels after several units had photobleached.

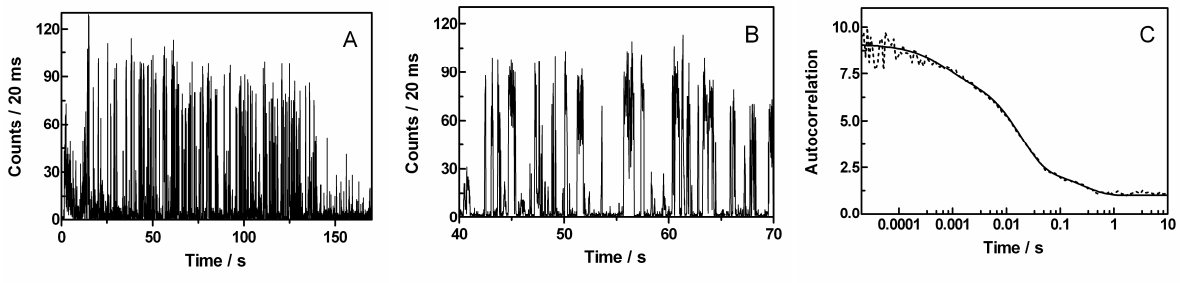

Figure 7.7 **A single molecule trajectory of mRFP1** (A) A single-molecule fluorescence trajectory of mRFP1 immobilized in PVA matrix. (B) Zoom in A) from 40-70 s. (C) Typical autocorrelation curve of an immobilized single-molecule and fit with a 3-exponential function (540 μs, 1.8 ms, 250 ms).

Autocorrelation curves of the fluorescence intensity of an immobilized molecule of mRFP1 could be typically fitted with 3 exponentials of timescales of hundreds of μs, a few ms, and hundreds of ms. Figure 7.7 *C* shows an example with autocorrelation times of 540 μs, 1.8 ms and 250 ms. The fastest of these processes might represent the flickering observed in the ensemble FCS measurements. The relaxation time found in the ensemble FCS measurements typically was <100 μs, but the difference might be attributed to the higher excitation intensity used in the ensemble FCS experiments. The slower processes might reflect protein conformational changes in a slower timescale (see Discussion). The distribution of

fluorescence lifetimes, analyzed in bins of 2000 photons, is quite broad, but peaks at about the main value found in bulk experiments (1.8 ns, see next section).

7.3.3. Time-resolved fluorescence spectroscopy of mRFP1 and mFruits

We performed time-resolved fluorescence measurements on bulk solutions of mRFPs to gain more insight in their excited-state dynamics. mRFP1 has main fluorescence decay of 1.8 ns (96%) and a small component of ~0.4 ns (4%) when excited at 543 nm and monitored at 610 nm (Figure 7.8 left). A non-pure monoexponential decay for mRFP1 has been already reported in literature (Stepanenko et al., 2004). mStrawberry (Figure 7.8 middle) and mCherry (Figure 7.8 right) also decayed biexponentially, the former with components of 2.1 ns (85%) and ~1 ns (15%) and the latter with 1.6 ns (90%) and ~0.7 ns (10%). The presence of multiexponential decays in GFPs could be associated with conformational changes (see Discussion). In contrast, DsRed decayed monoexponentially when excited at 543 nm (Heikal et al., 2000; Cotlet et al., 2001a). A multiexponential decay was found when excited at 470 nm due to energy transfer processes from the immature subunits of the tetramer, but this process cannot occur in the case of the monomer.

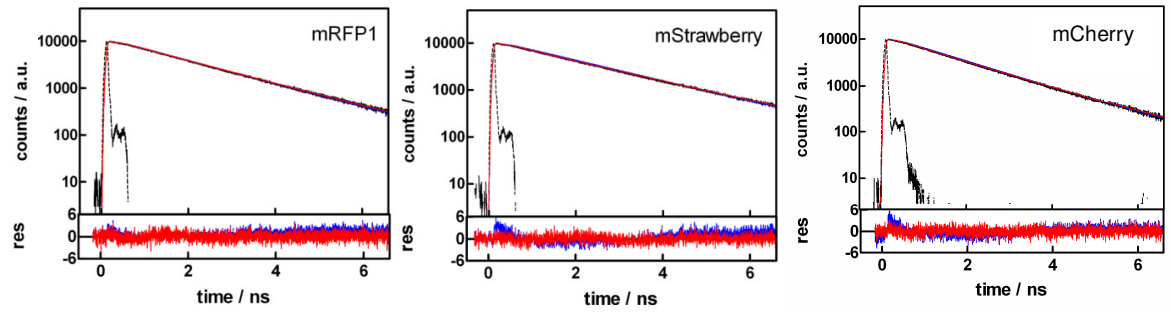

Figure 7.8 **Fluorescence decays of the mRFPs** – Fluorescence decay of (left) mRFP1, (middle) mStrawberry and (right) mCherry (λdet = 610 nm, solid line) and instrument response function (dotted line). Instrument response function (dotted line). Fit and residuals of a 1-exponential function (blue) and a 2-exponential function (red).

7.3.4. Structural basis for the pH-dependence of the mRFPs

It has been recently pointed out that Glu-215 is the key residue responsible for the pH-induced spectral shift of mCherry and mStrawberry, as it is the only residue in the vicinity of the chromophore that markedly changes its conformation upon an increase of the pH from 9.5 to 10.5 (Figure 7.9 top) (Shu et al., 2006). Although such a pKa is very high for a Glu (pKa for the γ-carboxyl group of Glu in solution is ~4.3), there are reports in the literature

for such high pKa values in some proteins (Forsyth et al., 2002) when the acid form is strongly stabilized. It is very intriguing that only the mRFPs show a pH-dependent red-shift and flickering, whereas all RFPs contain the same Glu-215 residue.

Figure 7.9 *A and B* show a stereo image of a 5 Å sphere around Glu-215 in mStrawberry and DsRed, respectively. Table 7.1 shows the residues surrounding Glu-215 that differ between DsRed and the mRFPs. The side chain of residue 41 is probably too distant from the side chain of Glu-215 to have an effect. Residues 42 and 44 are very similar among all RFPs. Residues 197 and 217 on the other hand differ strongly. While these residues are hydrophobic in the mRFPs, they are polar and involved in H-bond formation in DsRed. From the crystal structure of DsRed it is clear that Glu-215 forms a salt bridge with Lys-70, that is further stabilized by H-bridges with Ser-197 and Thr-217 (Yarbrough et al., 2001). The deprotonated form of Glu-215 is thus strongly stabilized through the salt bridge, causing a very low pKa of the carboxylic acid and providing an explanation for the lack of a pH-dependence of the fluorescence in DsRed.

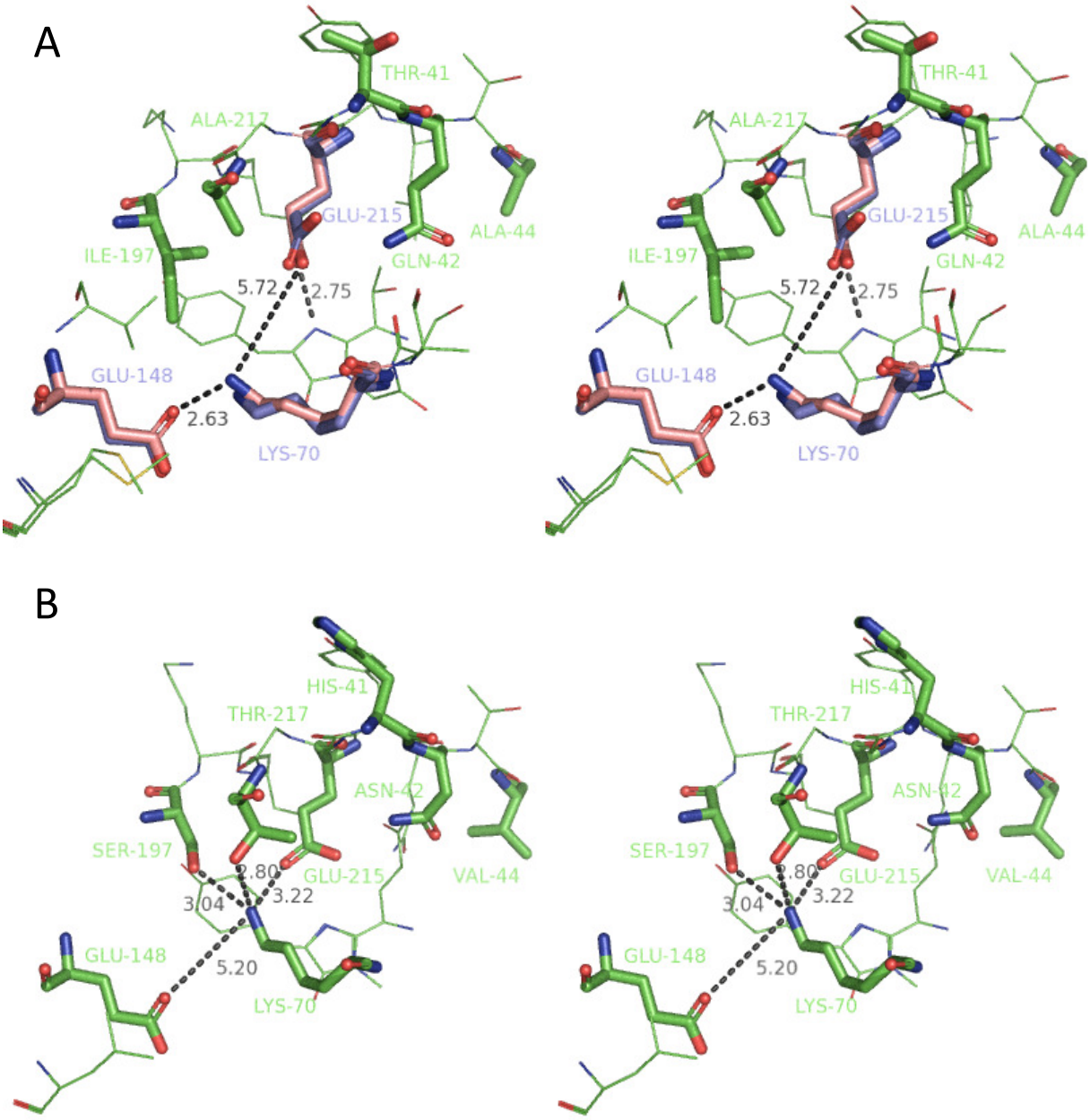

Figure 7.9 **Ball and stick stereo representations of the 5 Å-radius environment of Glu-215 in (A) mStrawberry and (B) DsRed** – Residues that differ between mStrawberry and DsRed or have a markedly different conformation in acid or base are explicitly depicted as sticks. In mStrawberry Lys-70, Glu-148 and Glu-215 are depicted pink in acid and blue in conjugated base conformation. H-bonds are depicted in gray with the distance in Å. The figures were made with the PyMOL Molecular Graphics System (2002, DeLano Scientific, Palo Alto, CA, USA).

Table 7.1 **Residues surrounding Glu-215 in RFPs that differ between DsRed and the mRFPs**

Position	DsRed	mRFPs
41	His	Thr
42	Asn	Gln
44	Val	Ala
197	**Ser**	**Ile**
217	**Thr**	**Ala**

In mStrawberry, however, the distance between the positively charged Lys-70 and the carboxylic acid oxygen of Glu-215 is increased to 5.72 Å and the local environment becomes more hydrophobic due to the presence of the two hydrophobic residues Ile-197 and Ala-217. Lys-70 thus forms a salt bridge with Glu-148 and moreover a H-bond is formed between the protonated Glu-217 and the imidazolinone nitrogen of the chromophore. This explains the high pKa of Glu-215. By way of comparison, in eqFP611, residue 197 is a His and 217 is an Ala. From the crystal structure, it is apparent that there is a H-bond between His-197 and Glu-215, stabilizing the carboxylate form and explaining the lack of pH-dependence (Malvezzi-Campeggi et al., 2001; Wiedenmann et al., 2002). Our observed pH-dependence can now be explained in terms of this structural information. At neutral pH, a H-bond between the chromophore and Glu-215 exists, which pulls the chromophore out of its planar conformation (Shu et al., 2006) causing a red shift of the spectrum and a decreased brightness. At a pH above the (high) pKa of Glu-215, the H-bond between the latter and the chromophore is broken, and as a result the chromophore will become more planar and thus brighter. Due to the breakage of the H-bond, conformational rearrangements will be favored, and a more frequent dark state formation occurs. A similar pH-dependence has been recently observed in a related chromoprotein (Battad et al., 2007).

The transition-pH for the mRFPs is also around the pKa of an ε-amino group of a Lys (10.5 in solution). However, in the acid and basic structures of the mRFPs Lys-70 is forming a salt bridge with Glu-148, such that its deprotonation as a cause of the transition can be excluded. This interpretation is confirmed by the fact that Lys-70 in mStrawberry has only a slightly different conformation at pH 10.5 as compared to pH 9.5.

7.4. Discussion

7.4.1. Dark states of mRFPs

mRFP1, mCherry and mStrawberry all reveal a pH- and I_{exc}-dependent flickering, as shown in this work by means of FCS. As mentioned in the introduction, light-induced flickering has already been observed before with FCS and single-molecule spectroscopy in RFPs such as DsRed and eqFP611 (Malvezzi-Campeggi et al., 2001; Wiedenmann et al., 2002; Schenk et al., 2004a). The conformation of the chromophore can have a large influence on the spectroscopic properties of a protein. Both *cis*- (DsRed and variants hereof) and *trans*- (eqFP611) conformations around the methylene bridge between the two cyclic parts can be fluorescent. In addition, coplanarity of the hydroxyphenyl and the imidazolinone moieties of the chromophore seems to be indispensable for fluorescence (Prescott et al., 2003; Henderson and Remington, 2006). It was suggested that conformational rearrangements of the chromophore might be responsible for flickering in these red FPs, since the timescales are consistent with these changes (Schenk et al., 2004a). On the other hand, amino acids in close proximity of the chromophore can also affect the fluorescence of the latter. The pH-dependence of the mRFP1 fluorescence, as evidenced here, is most probably a consequence of an altered chromophore environment (see next section). Moreover, the looser H-bonding network in mRFP1 and mFruits as compared to DsRed might contribute to the conformational freedom of the chromophore in the β-barrel (Campbell et al., 2002).

Our results for the time-resolved fluorescence experiments are consistent with conformational rearrangements of the chromophore in mRFP1 and mFruits. Similar biexponential decays in other red FPs such as eqFP611 and HcRed have been associated with the presence of two different conformers (Loos et al., 2006; Cotlet et al., 2006). Furthermore, we showed here with FCS that the conformers interconvert in the excited state, since the flickering is light-induced. In the case of DsRed, the pure monoexponential decay is a consequence of the reduced conformational mobility in the excited-state.

As for the blinking of immobilized mRFP1, the fact that the flickering is slowed down in the rigid PVA matrix would be consistent with the attribution of the process to a conformational change (Cotlet et al., 2006). It was previously shown that binding of eqFP611 to polyethylene glycol-covered glass surfaces did not affect the flickering rates as compared to solution (Schenk et al., 2004a), although the latter process was suppressed when eqFP611 was deposited on bare glass. The PVA matrix used in our single-molecule experiments might be

slightly more rigid than polyethylene glycol, but still allows a certain conformational freedom. The slight effect of solvent viscosity on the flickering component observed in FCS is also consistent with the above (although viscosity will clearly affect more the diffusion component). On the other hand, autocorrelation times in the tens of ms timescale have been previously assigned to the interconversion between *cis* and *trans* conformations mediated by polymer motions in immobilized HcRed (Cotlet et al., 2006). A hybrid quantum mechanics/molecular dynamics study might be useful to reveal the mechanisms of *cis-trans* isomerization and/or loss of chromophore planarity that are responsible for the observed fluorescence flickering in mRFPs, as was done previously for GFP (Weber et al., 1999).

By comparing DsRed and eqFP611, the latter which has been shown to have a smaller oligomerization tendency than the former, it was already shown by Schenk et al. that oligomerization does not influence to a great extent the flickering dynamics observed with FCS, although the difference in oligomerization state between the two proteins in their experimental conditions is somewhat inconclusive. We have found with single-molecule FCS (see section 'Single-molecule spectroscopy of immobilized mRFP1') that both immobilized mRFP1 and DsRed showed flickering, although for DsRed the flickering is somewhat more frequent in the lower intensity levels, after several units have photobleached. This situation would resemble that of mRFP1, although we cannot rule out the effect of the photobleached units of DsRed acting as 'traps'.

7.4.2. Implications for cellular measurements

To illustrate this, we made a fusion protein of mRFP and eGFP, expressed it inside live human HeLaP4 cells and performed FCS measurements, as described in (De Rijck et al., 2006).

Since the fluorescence of the fused protein can be monitored in the green and the red channel simultaneously (with two-colour excitation and detection), FCS analysis of both mRFP1- and eGFP-fluorescence is possible. Table 7.2 summarizes the results after fitting the two autocorrelation curves to a model with one exponential and two diffusion components (the fast representing free diffusion and the slow representing hindered diffusion). It is clear that fitting of the mRFP1 autocorrelation curve gives a non-realistic diffusion time for the freely diffusing protein (i.e., 39 μs for mRFPs vs 380 μs for eGFP), despite the fact that the flickering process is represented in the fitting. If the protein complex is bigger and diffusion is thus slowed down, then the processes can be separated well, as was the case in our *in vitro*

viscosity measurements. On the other hand, the sensitivity of dark state formation on the excitation intensity could also be exploited to separate diffusion and dark states. This would simplify the use of RFPs in cellular measurements considerably, but requires the RFP to be compatible with higher excitation intensities than typically used in cells.

Table 7.2 **Fit results for intracellular FCS on mRFP-eGFP** – Fitting was performed using a 2-component diffusion model with one exponential. Goodness-of-fit was judged by looking both at the residual $\delta G(\tau)$ curve and the χ^2-value.

FCS channel		τ (µs)	Fraction (%)
eGFP	Free diffusion	380 ± 70	74.8 ± 13.6
	Hindered diffusion	16405 ± 2220	25.2 ± 13.6
	Photophysics (triplet)	4 ± 2	16 ± 2
mRFP1	Free diffusion	39 ± 17	65.5 ± 8.0
	Hindered diffusion	2481 ± 840	34.5 ± 8.0
	Photophysics (flickering)	63 ± 11	43 ± 7

7.5. Conclusion

Fluorescence flickering in (m)RFPs seems to be a general behavior that might impair the analysis in FCS and FRET experiments. Unveiling the mechanism(s) responsible for this light-induced process may assist in the development of new and better mutants. In this work, we focused on the presence of this flickering in the monomeric RFPs mRFP1, mCherry and mStrawberry and provide insight in the pH-dependence of the fluorescence flickering. Taking into account previous results on other RFPs, we offer a global picture of the structural basis of fluorescence flickering in (monomeric) RFPs. Further evolution of mRFPs should specifically address this problem, since we show that the implications in cellular measurements can be serious.

Part IV – General conclusions

Identifying and characterizing intracellular interactions of a protein is typically performed by setting up a robust biochemical assay that specifically detects interacting partners in the presence of the protein. As with all research techniques, biochemical assays have their strengths and limitations. For example, although these assays do allow for the identification and characterization of cellular proteins, the biological content and three-dimensional architecture of the cell and its subcompartments creates an environment that can hardly be imitated in any *ex vivo* or *in vitro* assay. Secondly, in many interaction assays such as chromatin binding assays, pull-down and co-immunoprecipitation the time dimension is lost, only a static non-equilibrium picture of protein behavior and properties is obtained, whereas *in vivo* protein interactions occur dynamically.

Fluorescence correlation (FCS) and cross-correlation (FCCS) spectroscopy are two classic examples of biophysical research techniques that can be used to provide information at spatial and temporal scales that no single biochemical technique can approach. These techniques allow to measure the concentration, mobility and interactions of proteins inside a living functional cell, which opens a new dimension in studying protein function.

The rationale of this doctoral research project has been to gain insight into the intracellular function of a human transcriptional co-activator, LEDGF/p75, in view of its known importance for the human immunodeficiency virus, with the specific aim of investigating this protein in the time dimension. LEDGF/p75 has been suggested to guide integration towards active regions in the genome, through its known strong interaction with HIV-1 integrase, the viral enzyme that catalyzes the insertion of the viral genome into that of the infected immune cell. How does its strong affinity for chromatin, as suggested from in vitro experiments, relate to this?

We have clearly shown with intracellular FCS that LEDGF/p75 interacts extraordinarily dynamically with chromatin, mainly through non-specific interactions. We showed the

kinetics of association to be diffusion controlled with a high rate constant for dissociation (k_{off} > 38.6 s^{-1}) and a high concentration of (aspecific) binding sites. LEDGF/p75 thus continuously scans the genome, and finding a specific binding site amenable to transcriptional activation is thus likely a matter of chance. We are the first to demonstrate ultrafast chromatin binding kinetics for a transcription factor.

This dynamic profile changed drastically when HIV-1 integrase is co-expressed in cells. We showed order-of-magnitude slower dynamics of the LEDGF/p75-IN complex and pin-pointed these slow dynamics to the PWWP domain of LEDGF/p75, a known DNA/chromatin interacting domain. We finally investigated the stoichiometry of the complex and determined that IN can form oligomers independent of the presence of LEDGF/p75.

We have set up a set of intracellular assays to probe for the presence or absence of the LEDGF/p75-IN interaction, based on FCCS. Due to a higher order stoichiometry, the interacting complex likely contains multiple LEDGF/p75 and IN partners, we had to target the protein complex away from chromatin in order to quantify the interaction with FCCS. By quantitative analysis, we could calculate for the first time an intracellular dissociation constant of LEDGF/p75 and IN of around 400 nM, implying a fairly strong intracellular affinity. In the nucleus, competition with other co-factors of LEDGF/p75 and IN, a weaker apparent affinity was observed. The slow dynamics of the LEDGF/p75-IN complex may be advantageous for the viral integration complex, because it confers more time to IN to perform proper integration in the host cell genome. Future work will consist in investigating the contribution of the slow dynamics to viral integration efficiency. We speculate that the mechanism of the transcriptional co-activator, i.e. aspecific genome scanning followed by arrest at defined sites of gene activation, probably through association with the transcription machinery, has been hijacked by the lentiviral integration machinery to guarantee both high frequency of integration sites and proviral gene activation after integration. Future work on HIV will also consist in the quantification of the affinity of IN in cells for other binding partners, that could possibly take over the role of LEDGF/p75 when its interaction with IN is targeted by inhibitors. The relationship in terms of chromatin and protein binding affinity of binding partners, IN and LEDGF/p75 will be further investigated. Furthermore, chromatin binding of LEDGF/p75 will be investigated in view of its known roles in oncogenesis and autoimmunity. The importance of chromatin binding affinity for the

complete viral pre-integration complex will be analysed by studying the dynamics of integration complexes in cells depleted for LEDGF/p75.

Throughout the doctoral research, an important goal has been to verify the performance FCS and FCCS for measurements in living cells. The FCS technique was carefully investigated in a controlled *in vitro* environment. We verified the effects of the optical density, cover glass thickness, temperature, confocal pinhole, optical saturation and detector afterpulsing and provided an adequate method to calibrate a two-color setup. From these measurements we conclude that FCS can indeed work well in living cells, provided that are the above mentioned experimental parameters are carefully considered. We also studied the complex photophysics of monomeric red fluorescent proteins in detail with different approaches and showed that light-enhanced dark state formation is related to intra-protein chromophore mobility and can sometimes make quantitative FCS analysis difficult.

Finally, the performance of the FCCS technique for measuring interactions in cells by means of double fluorescent protein fusion, was investigated by quantitative analysis of a control protein fusion inside living cells. From these experiment, we had to conclude that the red fluorescent protein partner, be it mRFP1 or mCherry, suffers from low photostability, which renders a vast percentage of the population photobleached, with a negative effect on the measurement as a result. We have also experimentally shown that cross-talk of the green fluorochrome into the red detection channel and interfluorochrome FRET has a large effect on the measurement, and conclude that it is of crucial importance to quantify these effects for each protein-protein interaction to be analysed, if FCCS is to be used to calculate intracellular binding constants. Fluorescent protein cross-correlation spectroscopy has the potential of becoming the biophysical tool of choice for quantifying protein-protein interactions in cells, because the absolute concentrations of both partners and their complex can directly be measured. Future work will consist in applying pulsed-interleaved excitation or temporal laser switching FCCS on current and new fluorescent proteins. The photophysics of red fluorescent proteins and the search for a better fluorescence couple for FCCS will be continued, since the method has the potential of becoming an artefact-free quantitative technique for measuring protein-protein interactions in cells and can be used, in terms of drug-discovery, for the intracellular validation of inhibitors.

List of References

1. Ahuja, H. G., Hong, J. M., Aplan, P. D., Tcheurekdjian, L., Forman, S. J., and Slovak, M. L. (2000). t(9;11)(p22;p15) in acute myeloid leukemia results in a fusion between NUP98 and the gene encoding transcriptional coactivators p52 and p75-lens epithelium-derived growth factor (LEDGF). Cancer Research *60*, 6227-6229.
2. Axelrod, D., Koppel, D. E., Schlessinger, J., Elson, E., and Webb, W. W. (1976). Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. Biophys. J. *16*, 1055-1069.
3. Bacia, K., Kim, S. A., and Schwille, P. (2006). Fluorescence cross-correlation spectroscopy in living cells. Nat. Methods *3*, 83-89.
4. Bacia, K., Majoul, I. V., and Schwille, P. (2002). Probing the endocytic pathway in live cells using dual-color fluorescence cross-correlation analysis. Biophys. J. *83*, 1184-1193.
5. Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., xler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W., and Montagnier, L. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science *220*, 868-871.
6. Bartholomeeusen, K., Christ, F., Hendrix, J., Rain, J. C., Emiliani, S., Benarous, R., Debyser, Z., Gijsbers, R., and De Rijck, J. (2009). Lens Epithelium Derived Growth Factor/p75 interacts with the transposase derived DDE domain of pogZ. J. Biol. Chem. *284*, 11467-11477.
7. Bartholomeeusen, K., De Rijck, J., Busschots, K., Desender, L., Gijsbers, R., Emiliani, S., Benarous, R., Debyser, Z., and Christ, F. (2007). Differential interaction of HIV-1 integrase and JPO2 with the C terminus of LEDGF/p75. J. Mol. Biol. *372*, 407-421.
8. Battad, J. M., Wilmann, P. G., Olsen, S., Byres, E., Smith, S. C., Dove, S. G., Turcic, K. N., Devenish, R. J., Rossjohn, J., and Prescott, M. (2007). A structural basis for the pH-dependent increase in fluorescence efficiency of chromoproteins. J. Mol. Biol. *368*, 998-1010.
9. Baudendistel, N., Muller, G., Waldeck, W., Angel, P., and Langowski, J. (2005). Two-hybrid fluorescence cross-correlation spectroscopy detects protein-protein interactions in vivo. Chemphyschem. *6*, 984-990.
10. Bereiter-Hahn, J., Fox, C. H., and Thorell, B. (1979). Quantitative reflection contrast microscopy of living cells. J. Cell Biol. *82*, 767-779.
11. Berland, K. M., So, P. T., Chen, Y., Mantulin, W. W., and Gratton, E. (1996). Scanning two-photon fluctuation correlation spectroscopy: particle counting measurements for detection of molecular aggregation. Biophys. J *71*, 410-420.
12. Berry, C., Hannenhalli, S., Leipzig, J., and Bushman, F. D. (2006). Selection of target sites for mobile DNA integration in the human genome. PLoS. Comput. Biol. *2*, e157.
13. Bohmer, M., Wahl, M., Rahn, H. J., Erdmann, R., and Enderlein, J. (2002). Time-resolved fluorescence correlation spectroscopy. Chemical Physics Letters *353*, 439-445.
14. Braeckmans, K., Peeters, L., Sanders, N. N., De Smedt, S. C., and Demeester, J. (2003). Three-dimensional fluorescence recovery after photobleaching with the confocal scanning laser microscope. Biophys. J. *85*, 2240-2252.

15. Briers, Y., Schmelcher, M., Loessner, M. J., Hendrix, J., Engelborghs, Y., Volckaert, G., and Lavigne, R. (2009). The high-affinity peptidoglycan binding domain of Pseudomonas phage endolysin KZ144. Biochem. Biophys. Res. Commun. *383*, 187-191.
16. Briggs, J. A., Wilk, T., Welker, R., Krausslich, H. G., and Fuller, S. D. (2003). Structural organization of authentic, mature HIV-1 virions and cores. EMBO J *22*, 1707-1715.
17. Brunsting, A. and Mullaney, P. F. (1974). Differential light scattering from spherical mammalian cells. Biophys. J. *14*, 439-453.
18. Burkhardt, M., Heinze, K. G., and Schwille, P. (2005). Four-color fluorescence correlation spectroscopy realized in a grating-based detection platform. Opt. Lett. *30*, 2266-2268.
19. Busschots, K., Vercammen, J., Emiliani, S., Benarous, R., Engelborghs, Y., Christ, F., and Debyser, Z. (2005). The interaction of LEDGF/p75 with integrase is lentiviral-specific and promotes DNA binding. J. Biol. Chem. *280*, 17841-17847.
20. Busschots, K., Voet, A., De, M. M., Rain, J. C., Emiliani, S., Benarous, R., Desender, L., Debyser, Z., and Christ, F. (2007). Identification of the LEDGF/p75 binding site in HIV-1 integrase. J. Mol. Biol. *365*, 1480-1492.
21. Cai, M., Zheng, R., Caffrey, M., Craigie, R., Clore, G. M., and Gronenborn, A. M. (1997). Solution structure of the N-terminal zinc binding domain of HIV-1 integrase. Nat. Struct. Biol. *4*, 567-577.
22. Campbell, R. E., Tour, O., Palmer, A. E., Steinbach, P. A., Baird, G. S., Zacharias, D. A., and Tsien, R. Y. (2002). A monomeric red fluorescent protein. Proc. Natl. Acad. Sci. U. S A *99*, 7877-7882.
23. Chen, J. C., Krucinski, J., Miercke, L. J., Finer-Moore, J. S., Tang, A. H., Leavitt, A. D., and Stroud, R. M. (2000). Crystal structure of the HIV-1 integrase catalytic core and C-terminal domains: a model for viral DNA binding. Proc. Natl. Acad. Sci. U. S. A *97*, 8233-8238.
24. Chen, K., Ou, X. M., Chen, G., Choi, S. H., and Shih, J. C. (2005a). R1, a novel repressor of the human monoamine oxidase A. J. Biol. Chem. *280*, 11552-11559.
25. Chen, Y., Muller, J. D., Ruan, Q., and Gratton, E. (2002). Molecular brightness characterization of EGFP in vivo by fluorescence fluctuation spectroscopy. Biophys. J. *82*, 133-144.
26. Chen, Y., Muller, J. D., So, P. T., and Gratton, E. (1999). The photon counting histogram in fluorescence fluctuation spectroscopy. Biophys. J. *77*, 553-567.
27. Chen, Y., Tekmen, M., Hillesheim, L., Skinner, J., Wu, B., and Muller, J. D. (2005b). Dual-color photon-counting histogram. Biophys. J. *88*, 2177-2192.
28. Cherepanov, P. (2007). LEDGF/p75 interacts with divergent lentiviral integrases and modulates their enzymatic activity in vitro. Nucleic Acids Res. *35*, 113-124.
29. Cherepanov, P., Ambrosio, A. L., Rahman, S., Ellenberger, T., and Engelman, A. (2005a). Structural basis for the recognition between HIV-1 integrase and transcriptional coactivator p75. Proc. Natl. Acad. Sci. U. S. A *102*, 17308-17313.
30. Cherepanov, P., Devroe, E., Silver, P. A., and Engelman, A. (2004). Identification of an evolutionarily conserved domain in human lens epithelium-derived growth factor/transcriptional co-activator p75 (LEDGF/p75) that binds HIV-1 integrase. J. Biol. Chem. *279*, 48883-48892.
31. Cherepanov, P., Maertens, G., Proost, P., Devreese, B., Van Beeumen, J., Engelborghs, Y., De Clercq, E., and Debyser, Z. (2003). HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. J. Biol. Chem. *278*, 372-381.
32. Cherepanov, P., Pluymers, W., Claeys, A., Proost, P., De, C. E., and Debyser, Z. (2000). High-level expression of active HIV-1 integrase from a synthetic gene in human cells. FASEB J. *14*, 1389-1399.

33. Cherepanov, P., Sun, Z. Y., Rahman, S., Maertens, G., Wagner, G., and Engelman, A. (2005b). Solution structure of the HIV-1 integrase-binding domain in LEDGF/p75. Nat. Struct. Mol. Biol. *12*, 526-532.
34. Christ, F., Thys, W., De Rijck, J., Gijsbers, R., Albanese, A., Arosio, D., Emiliani, S., Rain, J. C., Benarous, R., Cereseto, A., and Debyser, Z. (2008). Transportin-SR2 imports HIV into the nucleus. Curr. Biol. *18*, 1192-1202.
35. Ciuffi, A., Llano, M., Poeschla, E., Hoffmann, C., Leipzig, J., Shinn, P., Ecker, J. R., and Bushman, F. (2005). A role for LEDGF/p75 in targeting HIV DNA integration. Nat. Med. *11*, 1287-1289.
36. Cotlet, M., Habuchi, S., Whitier, J. E., Werner, J. H., De Schryver, F. C., Hofkens, J., and Goodwin, P. M. (2006). Single Molecule Spectroscopic Characterization of a Far-Red Fluorescent protein (HcRed) from the Anthozoa Coral *Heteractis Crispa*. Proc. SPIE *6098*, 609804-609811.
37. Cotlet, M., Hofkens, J., Habuchi, S., Dirix, G., Van Guyse, M., Michiels, J., Vanderleyden, J., and De Schryver, F. C. (2001a). Identification of different emitting species in the red fluorescent protein DsRed by means of ensemble and single-molecule spectroscopy. Proc. Natl. Acad. Sci. U. S. A *98*, 14398-14403.
38. Cotlet, M., Hofkens, J., Kohn, F., Michiels, J., Dirix, G., Van Guyse, M., Vanderleyden, J., and De Schryver, F. C. (2001b). Collective effects in individual oligomers of the red fluorescent coral protein DsRed. Chem. Phys. Lett. *336*, 415-423.
39. Craigie, R. (2001). HIV integrase, a brief overview from chemistry to therapeutics. J. Biol. Chem. *276*, 23213-23216.
40. Crank,J. (1975). The Mathematics of Diffusion. (Oxford: Clarendon Press).
41. Curl, C. L., Bellair, C. J., Harris, T., Allman, B. E., Harris, P. J., Stewart, A. G., Roberts, A., Nugent, K. A., and Delbridge, L. M. (2005). Refractive index measurement in viable cells using quantitative phase-amplitude microscopy and confocal microscopy. Cytometry A *65*, 88-92.
42. Daniels, T., Zhang, J. Y., Gutierrez, I., Elliot, M. L., Yamada, B., Heeb, M. J., Sheets, S. M., Wu, X. W., and Casiano, C. A. (2005). Antinuclear autoantibodies in prostate cancer: Immunity to LEDGF/p75, a survival protein highly expressed in prostate tumors and cleaved during apoptosis. Prostate *62*, 14-26.
43. Daugaard, M., Kirkegaard-Sorensen, T., Ostenfeld, M. S., Aaboe, M., Hoyer-Hansen, M., Orntoft, T. F., Rohde, M., and Jaattela, M. (2007). Lens epithelium-derived growth factor is an Hsp70-2 regulated guardian of lysosomal stability in human cancer. Cancer Res. *67*, 2559-2567.
44. De Rijck J., Vandekerckhove, L., Gijsbers, R., Hombrouck, A., Hendrix, J., Vercammen, J., Engelborghs, Y., Christ, F., and Debyser, Z. (2006a). Overexpression of the lens epithelium-derived growth factor/p75 integrase binding domain inhibits human immunodeficiency virus replication. J. Virol. *80*, 11498-11509.
45. De Rijck J., Vandekerckhove, L., Gijsbers, R., Hombrouck, A., Hendrix, J., Vercammen, J., Engelborghs, Y., Christ, F., and Debyser, Z. (2006b). Overexpression of the lens epithelium-derived growth factor/p75 integrase binding domain inhibits human immunodeficiency virus replication. J. Virol. *80*, 11498-11509.
46. De Rijck, J., Vandekerckhove, L., Gijsbers, R., Hombrouck, A., Hendrix, J., Vercammen, J., Engelborghs, Y., Christ, F., and Debyser, Z. (2006). Overexpression of the lens epithelium-derived growth factor/p75 integrase binding domain inhibits human immunodeficiency virus replication. J. Virol. *80*, 11498-11509.
47. Delelis, O., Carayon, K., Guiot, E., Leh, H., Tauc, P., Brochon, J. C., Mouscadet, J. F., and Deprez, E. (2008). Insight into the integrase-DNA recognition mechanism. A specific DNA-binding mode revealed by an enzymatically labeled integrase. J. Biol. Chem. *283*, 27838-27849.

48. Delelis, O., Parissi, V., Leh, H., Mbemba, G., Petit, C., Sonigo, P., Deprez, E., and Mouscadet, J. F. (2007). Efficient and specific internal cleavage of a retroviral palindromic DNA sequence by tetrameric HIV-1 integrase. PLoS. ONE. *2*, e608.
49. Deprez, E., Tauc, P., Leh, H., Mouscadet, J. F., Auclair, C., and Brochon, J. C. (2000). Oligomeric states of the HIV-1 integrase as measured by time-resolved fluorescence anisotropy. Biochemistry *39*, 9275-9284.
50. Dertinger, T., Colyer, R., Iyer, G., Weiss, S., and Enderlein, J. (2009). Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI). Proc. Natl. Acad. Sci. U. S. A *106*, 22287-22292.
51. Dertinger, T., Ewers, B., von der Hocht, I., and Enderlein, J. (2007a). Two-focus fluorescence correlation spectroscopy. Picoquant Application Notes 1-6.
52. Dertinger, T., Pacheco, V., von der Hocht, I., Hartmann, R., Gregor, I., and Enderlein, J. (2007b). Two-focus fluorescence correlation spectroscopy: a new tool for accurate and absolute diffusion measurements. Chemphyschem. *8*, 433-443.
53. Dietz, F., Franken, S., Yoshida, K., Nakamura, H., Kappler, J., and Gieselmann, V. (2002). The family of hepatoma-derived growth factor proteins: characterization of a new member HRP-4 and classification of its subfamilies. Biochemical Journal *366*, 491-500.
54. Digman, M. A., Dalal, R., Horwitz, A. F., and Gratton, E. (2008). Mapping the number of molecules and brightness in the laser scanning microscope. Biophys. J. *94*, 2320-2332.
55. Digman, M. A., Sengupta, P., Wiseman, P. W., Brown, C. M., Horwitz, A. R., and Gratton, E. (2005). Fluctuation correlation spectroscopy with a laser-scanning microscope: exploiting the hidden time structure. Biophys. J. *88*, L33-L36.
56. Digman, M. A., Wiseman, P. W., Choi, C., Horwitz, A. R., and Gratton, E. (2009a). Stoichiometry of molecular complexes at adhesions in living cells. Proceedings of the National Academy of Sciences *106*, 2170-2175.
57. Digman, M. A., Wiseman, P. W., Horwitz, A. R., and Gratton, E. (2009b). Detecting protein complexes in living cells from laser scanning confocal image sequences by the cross correlation raster image spectroscopy method. Biophys. J. *96*, 707-716.
58. Dyda, F., Hickman, A. B., Jenkins, T. M., Engelman, A., Craigie, R., and Davies, D. R. (1994). Crystal structure of the catalytic domain of HIV-1 integrase: similarity to other polynucleotidyl transferases. Science *266*, 1981-1986.
59. Eggeling, C., Ringemann, C., Medda, R., Schwarzmann, G., Sandhoff, K., Polyakova, S., Belov, V. N., Hein, B., von, M. C., Schonle, A., and Hell, S. W. (2009). Direct observation of the nanoscale dynamics of membrane lipids in a living cell. Nature *457*, 1159-1162.
60. Ehrenberg, M. and Rigler, R. (1974). Rotational brownian motion and fluorescence intensify fluctuations. Chemical Physics *4*, 390-401.
61. Eijkelenboom, A. P., Lutzke, R. A., Boelens, R., Plasterk, R. H., Kaptein, R., and Hard, K. (1995). The DNA-binding domain of HIV-1 integrase has an SH3-like fold. Nat. Struct. Biol. *2*, 807-810.
62. Eissenberg, J. C. and Elgin, S. C. (2000). The HP1 protein family: getting a grip on chromatin. Curr. Opin. Genet. Dev. *10*, 204-210.
63. Elson, E. L. and Magde, D. (1974). Fluorescence Correlation Spectroscopy .1. Conceptual Basis and Theory. Biopolymers *13*, 1-27.
64. Emiliani, S., Mousnier, A., Busschots, K., Maroun, M., Van, M. B., Tempe, D., Vandekerckhove, L., Moisant, F., Ben-Slama, L., Witvrouw, M., Christ, F., Rain, J. C., Dargemont, C., Debyser, Z., and Benarous, R. (2005). Integrase mutants defective for interaction with LEDGF/P75 are impaired in chromosome tethering and HIV-1 replication. J. Biol. Chem. *280*, 25517-25523.
65. Enderlein, J. and Gregor, I. (2005). Using fluorescence lifetime for discriminating detector afterpulsing in fluorescence-correlation spectroscopy. Review of Scientific Instruments *76*, 033102:1-5.

66. Enderlein, J., Gregor, I., Patra, D., Dertinger, T., and Kaupp, U. B. (2005). Performance of fluorescence correlation spectroscopy for measuring diffusion and concentration. Chemphyschem. *6*, 2324-2336.
67. Enderlein, J., Gregor, I., Patra, D., and Fitter, J. (2004). Art and artefacts of fluorescence correlation spectroscopy. Curr. Pharm. Biotechnol. *5*, 155-161.
68. Engelman, A. and Cherepanov, P. (2008). The lentiviral integrase binding protein LEDGF/p75 and HIV-1 replication. PLoS. Pathog. *4*, e1000046.
69. Farnet, C. M. and Bushman, F. D. (1997). HIV-1 cDNA integration: requirement of HMG I(Y) protein for function of preintegration complexes in vitro. Cell *88*, 483-492.
70. Fatma, N., Kubo, E., Chylack, L. T., Shinohara, T., Akagi, Y., and Singh, D. P. (2004). LEDGF regulation of alcohol and aldehyde dehydrogenases in lens epithelial cells: stimulation of retinoic acid production and protection from ethanol toxicity. American Journal of Physiology-Cell Physiology *287*, C508-C516.
71. Fatma, N., Singh, D. P., Shinohara, T., and Chylack, L. T. (2001). Transcriptional regulation of the antioxidant protein 2 gene, a thiol-specific antioxidant, by lens epithelium-derived growth factor to protect cells from oxidative stress. Journal of Biological Chemistry *276*, 48899-48907.
72. Faure, A., Calmels, C., Desjobert, C., Castroviejo, M., Caumont-Sarcos, A., Tarrago-Litvak, L., Litvak, S., and Parissi, V. (2005). HIV-1 integrase crosslinked oligomers are active in vitro. Nucleic Acids Res. *33*, 977-986.
73. Feder, T. J., Brust-Mascher, I., Slattery, J. P., Baird, B., and Webb, W. W. (1996). Constrained diffusion or immobile fraction on cell surfaces: a new interpretation. Biophys. J. *70*, 2767-2773.
74. Ferrand, P., Pianta, M., Kress, A., Aillaud, A., Rigneault, H., and Marguet, D. (2009). A versatile dual spot laser scanning confocal microscopy system for advanced fluorescence correlation spectroscopy analysis in living cell. Rev. Sci. Instrum. *80*, 083702.
75. Foldes-Papp, Z. (2005). How the molecule number is correctly quantified in two-color fluorescence cross-correlation spectroscopy: corrections for cross-talk and quenching in experiments. Curr. Pharm. Biotechnol. *6*, 437-444.
76. Forsyth, W. R., Antosiewicz, J. M., and Robertson, A. D. (2002). Empirical relationships between protein structure and carboxyl pKa values in proteins. Proteins *48*, 388-403.
77. Frankel, A. D. and Young, J. A. (1998). HIV-1: fifteen proteins and an RNA. Annu. Rev. Biochem. *67*, 1-25.
78. Gallay, P., Hope, T., Chin, D., and Trono, D. (1997). HIV-1 infection of nondividing cells through the recognition of integrase by the importin/karyopherin pathway. Proc. Natl. Acad. Sci. U. S. A *94*, 9825-9830.
79. Ganapathy, V. and Casiano, C. A. (2004). Autoimmunity to the nuclear autoantigen DFS70 (LEDGF): what exactly are the autoantibodies trying to tell us? Arthritis Rheum. *50*, 684-688.
80. Ganapathy, V., Daniels, T., and Casiano, C. A. (2003). LEDGE/p75: a novel nuclear autoantigen at the crossroads of cell survival and apoptosis. Autoimmunity Reviews *2*, 290-297.
81. Garcia-Parajo, M. F., Koopman, M., van Dijk, E. M., Subramaniam, V., and van Hulst, N. F. (2001). The nature of fluorescence emission in the red fluorescent protein DsRed, revealed by single-molecule detection. Proc. Natl. Acad. Sci. U. S. A *98*, 14392-14397.
82. Ge, H., Si, Y. Z., and Roeder, R. G. (1998a). Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. Embo Journal *17*, 6723-6729.
83. Ge, H., Si, Y. Z., and Wolffe, A. P. (1998b). A novel transcriptional coactivator, p52, functionally interacts with the essential splicing factor ASF/SF2. Molecular Cell *2*, 751-759.

84. Grand, F. H., Koduru, P., Cross, N. C., and Allen, S. L. (2005). NUP98-LEDGF fusion and t(9;11) in transformed chronic myeloid leukemia. Leuk. Res. *29*, 1469-1472.
85. Gregor, I., Patra, D., and Enderlein, J. (2005). Optical saturation in fluorescence correlation spectroscopy under continuous-wave and pulsed excitation. Chemphyschem. *6*, 164-170.
86. Gross, L. A., Baird, G. S., Hoffman, R. C., Baldridge, K. K., and Tsien, R. Y. (2000). The structure of the chromophore within DsRed, a red fluorescent protein from coral. Proc. Natl. Acad. Sci. U. S A *97*, 11990-11995.
87. Guiot, E., Carayon, K., Delelis, O., Simon, F., Tauc, P., Zubin, E., Gottikh, M., Mouscadet, J. F., Brochon, J. C., and Deprez, E. (2006). Relationship between the oligomeric status of HIV-1 integrase on DNA and enzymatic activity. J. Biol. Chem. *281*, 22707-22719.
88. Habuchi, S., Cotlet, M., Gensch, T., Bednarz, T., Haber-Pohlmeier, S., Rozenski, J., Dirix, G., Michiels, J., Vanderleyden, J., Heberle, J., De Schryver, F. C., and Hofkens, J. (2005). Evidence for the isomerization and decarboxylation in the photoconversion of the red fluorescent protein DsRed. J. Am. Chem. Soc. *127*, 8977-8984.
89. Hare, S., Di Nunzio, F., Labeja, A., Wang, J., Engelman, A., and Cherepanov, P. (2009a). Structural basis for functional tetramerization of lentiviral integrase. PLoS. Pathog. *5*, e1000515.
90. Hare, S., Shun, M. C., Gupta, S. S., Valkov, E., Engelman, A., and Cherepanov, P. (2009b). A novel co-crystal structure affords the design of gain-of-function lentiviral integrase mutants in the presence of modified PSIP1/LEDGF/p75. PLoS. Pathog. *5*, e1000259.
91. Haupts, U., Maiti, S., Schwille, P., and Webb, W. W. (1998). Dynamics of fluorescence fluctuations in green fluorescent protein observed by fluorescence correlation spectroscopy. Proc. Natl. Acad. Sci. U. S A *95*, 13573-13578.
92. Hayouka, Z., Rosenbluh, J., Levin, A., Loya, S., Lebendiker, M., Veprintsev, D., Kotler, M., Hizi, A., Loyter, A., and Friedler, A. (2007). Inhibiting HIV-1 integrase by shifting its oligomerization equilibrium. Proc. Natl. Acad. Sci. U. S. A *104*, 8316-8321.
93. Hebert, B., Costantino, S., and Wiseman, P. W. (2005). Spatiotemporal image correlation spectroscopy (STICS) theory, verification, and application to protein velocity mapping in living CHO cells. Biophys. J. *88*, 3601-3614.
94. Heikal, A. A., Hess, S. T., Baird, G. S., Tsien, R. Y., and Webb, W. W. (2000). Molecular spectroscopy and dynamics of intrinsically fluorescent proteins: coral red (dsRed) and yellow (Citrine). Proc. Natl. Acad. Sci. U. S A *97*, 11996-12001.
95. Heim, R., Cubitt, A. B., and Tsien, R. Y. (1995). Improved green fluorescence. Nature *373*, 663-664.
96. Hell, S. W. and Wichmann, J. (1994). Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. Opt. Lett. *19*, 780-782.
97. Henderson, J. N. and Remington, S. J. (2006). The kindling fluorescent protein: A transient photoswitchable marker. Physiology *21*, 162-170.
98. Hendrix, J., Flors, C., Dedecker, P., Hofkens, J., and Engelborghs, Y. (2008). Dark states in monomeric red fluorescent proteins studied by fluorescence correlation and single molecule spectroscopy. Biophys. J. *94*, 4103-4113.
99. Hess, S. T. and Webb, W. W. (2002). Focal volume optics and experimental artifacts in confocal fluorescence correlation spectroscopy. Biophys. J *83*, 2300-2317.
100. Hombrouck, A., De Rijck, J., Hendrix, J., Vandekerckhove, L., Voet, A., De Maeyer, M., Witvrouw, M., Engelborghs, Y., Christ, F., Gijsbers, R., and Debyser, Z. (2007). Virus Evolution Reveals an Exclusive Role for LEDGF/p75 in Chromosomal Tethering of HIV. PLoS. Pathog. *3*, e47.

101. Huang, A., Ho, C. S., Ponzielli, R., Barsyte-Lovejoy, D., Bouffet, E., Picard, D., Hawkins, C. E., and Penn, L. Z. (2005). Identification of a novel c-Myc protein interactor, JPO2, with transforming activity in medulloblastoma cells. Cancer Res. *65*, 5607-5619.
102. Huang, T. S., Myklebust, L. M., Kjarland, E., Gjertsen, B. T., Pendino, F., Bruserud, O., Doskeland, S. O., and Lillehaug, J. R. (2007). LEDGF/p75 has increased expression in blasts from chemotherapy-resistant human acute myelogenic leukemia patients and protects leukemia cells from apoptosis in vitro. Mol. Cancer *6*, 31.
103. Hughes, S., Jenkins, V., Dar, M. J., Engelman, A., and Cherepanov, P. (2010). Transcriptional co-activator LEDGF interacts with Cdc7-activator of S-phase kinase (ASK) and stimulates its enzymatic activity. J. Biol. Chem. *285*, 541-554.
104. Hussey, D. J., Moore, S., Nicola, M., and Dobrovic, A. (2001). Fusion of the NUP98 gene with the LEDGF/p52 gene defines a recurrent acute myeloid leukemia translocation. BMC. Genet. *2*, 20(1)-20(4).
105. Hwang, L. C., Gosch, M., Lasser, T., and Wohland, T. (2006). Simultaneous multicolor fluorescence cross-correlation spectroscopy to detect higher order molecular interactions using single wavelength laser excitation. Biophys. J. *91*, 715-727.
106. Hwang, L. C. and Wohland, T. (2004). Dual-color fluorescence cross-correlation spectroscopy using single laser wavelength excitation. Chemphyschem. *5*, 549-551.
107. Hwang, L. C. and Wohland, T. (2007). Recent advances in fluorescence cross-correlation spectroscopy. Cell Biochem. Biophys. *49*, 1-13.
108. Ikegame, K., Yamamoto, M., Kishima, Y., Enomoto, H., Yoshida, K., Suemura, M., Kishimoto, T., and Nakamura, H. (1999). A new member of a hepatoma-derived growth factor gene family can translocate to the nucleus. Biochem. Biophys. Res. Commun. *266*, 81-87.
109. Izumoto, Y., Kuroda, T., Harada, H., Kishimoto, T., and Nakamura, H. (1997). Hepatoma-derived growth factor belongs to a gene family in mice showing significant homology in the amino terminus. Biochem. Biophys. Res. Commun. *238*, 26-32.
110. Jach, G., Pesch, M., Richter, K., Frings, S., and Uhrig, J. F. (2006). An improved mRFP1 adds red to bimolecular fluorescence complementation. Nat. Methods *3*, 597-600.
111. Jenkins, T. M., Engelman, A., Ghirlando, R., and Craigie, R. (1996). A soluble active mutant of HIV-1 integrase: involvement of both the core and carboxyl-terminal domains in multimerization. J. Biol. Chem. *271*, 7712-7718.
112. Jenkins, T. M., Hickman, A. B., Dyda, F., Ghirlando, R., Davies, D. R., and Craigie, R. (1995). Catalytic domain of human immunodeficiency virus type 1 integrase: identification of a soluble mutant by systematic replacement of hydrophobic residues. Proc. Natl. Acad. Sci. U. S A *92*, 6057-6061.
113. Jimenez-Banzo, A., Nonell, S., Hofkens, J., and Flors, C. (2008). Singlet oxygen photosensitization by EGFP and its chromophore HBDI. Biophys. J. *94*, 168-172.
114. Jung, G., Wiehler, J., and Zumbusch, A. (2005). The photophysics of green fluorescent protein: Influence of the key amino acids at positions 65, 203, and 222. Biophys. J. *88*, 1932-1947.
115. Kalpana, G. V., Marmon, S., Wang, W., Crabtree, G. R., and Goff, S. P. (1994). Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. Science *266*, 2002-2006.
116. Kapusta, P., Wahl, M., Benda, A., Hof, M., and Enderlein, J. (2007). Fluorescence lifetime correlation spectroscopy. J. Fluoresc. *17*, 43-48.
117. Karki, R. G., Tang, Y., Burke, T. R., Jr., and Nicklaus, M. C. (2004). Model of full-length HIV-1 integrase complexed with viral DNA as template for anti-HIV drug design. J. Comput. Aided Mol. Des *18*, 739-760.

118. Kask, P., Palo, K., Ullmann, D., and Gall, K. (1999). Fluorescence-intensity distribution analysis and its application in biomolecular detection technology. Proc. Natl. Acad. Sci. U. S. A *96*, 13756-13761.
119. Kask, P., Piksarv, P., Mets, U., Pooga, M., and Lippmaa, E. (1987). Fluorescence correlation spectroscopy in the nanosecond time range: rotational diffusion of bovine carbonic anhydrase B. Eur. Biophys. J. *14*, 257-261.
120. Kim, S. A., Heinze, K. G., Bacia, K., Waxham, M. N., and Schwille, P. (2005). Two-photon cross-correlation analysis of intracellular reactions with variable stoichiometry. Biophys. J. *88*, 4319-4336.
121. Kim, S. A., Heinze, K. G., Waxham, M. N., and Schwille, P. (2004). Intracellular calmodulin availability accessed with two-photon cross-correlation. Proc. Natl. Acad. Sci. U. S. A *101*, 105-110.
122. Kirsch, R. D. and Joly, E. (1998). An improved PCR-mutagenesis strategy for two-site mutagenesis or sequence swapping between related genes. Nucleic Acids Res. *26*, 1848-1850.
123. Kogure, T., Karasawa, S., Araki, T., Saito, K., Kinjo, M., and Miyawaki, A. (2006). A fluorescent variant of a protein from the stony coral Montipora facilitates dual-color single-laser fluorescence cross-correlation spectroscopy. Nat. Biotechnol. *24*, 577-581.
124. Kohl, T., Haustein, E., and Schwille, P. (2005). Determining protease activity in vivo by fluorescence cross-correlation analysis. Biophys. J. *89*, 2770-2782.
125. Kolin, D. L., Ronis, D., and Wiseman, P. W. (2006). k-Space image correlation spectroscopy: a method for accurate transport measurements independent of fluorophore photophysics. Biophys. J. *91*, 3061-3075.
126. Lamb, D. C., Schenk, A., Rocker, C., Scalfi-Happ, C., and Nienhaus, G. U. (2000). Sensitivity enhancement in fluorescence correlation spectroscopy of multiple species using time-gated detection. Biophys. J. *79*, 1129-1138.
127. Lanni, F., Waggoner, A. S., and Taylor, D. L. (1985). Structural organization of interphase 3T3 fibroblasts studied by total internal reflection fluorescence microscopy. J. Cell Biol. *100*, 1091-1102.
128. Lee, M. S. and Craigie, R. (1998). A previously unidentified host protein protects retroviral DNA from autointegration. Proc. Natl. Acad. Sci. U. S A *95*, 1528-1533.
129. Lee, S. P., Xiao, J., Knutson, J. R., Lewis, M. S., and Han, M. K. (1997). Zn2+ promotes the self-association of human immunodeficiency virus type-1 integrase in vitro. Biochemistry *36*, 173-180.
130. Levene, M. J., Korlach, J., Turner, S. W., Foquet, M., Craighead, H. G., and Webb, W. W. (2003). Zero-mode waveguides for single-molecule analysis at high concentrations. Science *299*, 682-686.
131. Lewinski, M. K., Yamashita, M., Emerman, M., Ciuffi, A., Marshall, H., Crawford, G., Collins, F., Shinn, P., Leipzig, J., Hannenhalli, S., Berry, C. C., Ecker, J. R., and Bushman, F. D. (2006). Retroviral DNA integration: viral and cellular determinants of target-site selection. PLoS. Pathog. *2*, e60.
132. Lin, C. W. and Engelman, A. (2003). The barrier-to-autointegration factor is a component of functional human immunodeficiency virus type 1 preintegration complexes. J. Virol. *77*, 5030-5036.
133. Liu, P., Sudhaharan, T., Koh, R. M., Hwang, L. C., Ahmed, S., Maruyama, I. N., and Wohland, T. (2007). Investigation of the dimerization of proteins from the epidermal growth factor receptor family by single wavelength fluorescence cross-correlation spectroscopy. Biophys. J. *93*, 684-698.
134. Ljungquist, E., Khatoon, H., DuBow, M., Ambrosio, L., De, B. F., and Bukhari, A. I. (1979). Integration of bacteriophage mu DNA. Cold Spring Harb. Symp. Quant. Biol. *43 Pt 2*, 1151-1158.

135. Llano, M., Delgado, S., Vanegas, M., and Poeschla, E. M. (2004a). Lens epithelium-derived growth factor/p75 prevents proteasomal degradation of HIV-1 integrase. Journal of Biological Chemistry *279*, 55570-55577.
136. Llano, M., Saenz, D. T., Meehan, A., Wongthida, P., Peretz, M., Walker, W. H., Teo, W., and Poeschla, E. M. (2006a). An essential role for LEDGF/p75 in HIV integration. Science *314*, 461-464.
137. Llano, M., Vanegas, M., Fregoso, O., Saenz, D., Chung, S., Peretz, M., and Poeschla, E. M. (2004b). LEDGF/p75 determines cellular trafficking of diverse lentiviral but not murine oncoretroviral integrase proteins and is a component of functional lentiviral preintegration complexes. J. Virol. *78*, 9524-9537.
138. Llano, M., Vanegas, M., Hutchins, N., Thompson, D., Delgado, S., and Poeschla, E. M. (2006b). Identification and characterization of the chromatin-binding domains of the HIV-1 integrase interactor LEDGF/p75. J. Mol. Biol. *360*, 760-773.
139. Lodi, P. J., Ernst, J. A., Kuszewski, J., Hickman, A. B., Engelman, A., Craigie, R., Clore, G. M., and Gronenborn, A. M. (1995). Solution structure of the DNA binding domain of HIV-1 integrase. Biochemistry *34*, 9826-9833.
140. Loos, D. C., Habuchi, S., Flors, C., Hotta, J. I., Wiedenmann, J. R., Nienhaus, G. U., and Hofkens, J. (2006). Photoconversion in the red fluorescent protein from the sea anemone Entacmaea quadricolor: Is cis-trans isomerization involved? J. Am. Chem. Soc. *128*, 6270-6271.
141. Lukasik, S. M., Cierpicki, T., Borloz, M., Grembecka, J., Everett, A., and Bushweller, J. H. (2006). High resolution structure of the HDGF PWWP domain: a potential DNA binding domain. Protein Sci. *15*, 314-323.
142. Lutzke, R. A. and Plasterk, R. H. (1998). Structure-based mutational analysis of the C-terminal DNA-binding domain of human immunodeficiency virus type 1 integrase: critical residues for protein oligomerization and DNA binding. J. Virol. *72*, 4841-4848.
143. Maertens, G., Cherepanov, P., Debyser, Z., Engelborghs, Y., and Engelman, A. (2004). Identification and characterization of a functional nuclear localization signal in the HIV-1 integrase interactor LEDGF/p75. J. Biol. Chem. *279*, 33421-33429.
144. Maertens, G., Cherepanov, P., Pluymers, W., Busschots, K., De Clercq, E., Debyser, Z., and Engelborghs, Y. (2003). LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. J. Biol. Chem. *278*, 33528-33539.
145. Maertens, G., Vercammen, J., Debyser, Z., and Engelborghs, Y. (2005). Measuring protein-protein interactions inside living cells using single color fluorescence correlation spectroscopy. Application to human immunodeficiency virus type 1 integrase and LEDGF/p75. FASEB J. *19*, 1039-1041.
146. Maertens, G. N., Cherepanov, P., and Engelman, A. (2006). Transcriptional co-activator p75 binds and tethers the Myc-interacting protein JPO2 to chromatin. J. Cell Sci. *119*, 2563-2571.
147. Magde, D., Elson, E. L., and Webb, W. W. (1972). Fluorescence correlation spectroscopy. Biopolymers *13*, 29-61.
148. Malvezzi-Campeggi, F., Jahnz, M., Heinze, K. G., Dittrich, P., and Schwille, P. (2001). Light-induced flickering of DsRed provides evidence for distinct and interconvertible fluorescent states. Biophys. J. *81*, 1776-1785.
149. Margalit, A., Liu, J., Fridkin, A., Wilson, K. L., and Gruenbaum, Y. (2005). A lamin-dependent pathway that regulates nuclear organization, cell cycle progression and germ cell development. Novartis. Found. Symp. *264*, 231-240.
150. Marshall, H. M., Ronen, K., Berry, C., Llano, M., Sutherland, H., Saenz, D., Bickmore, W., Poeschla, E., and Bushman, F. D. (2007). Role of PSIP1/LEDGF/p75 in Lentiviral Infectivity and Integration Targeting. PLoS. ONE. *2*, e1340.

151. Marshall, W. F., Straight, A., Marko, J. F., Swedlow, J., Dernburg, A., Belmont, A., Murray, A. W., Agard, D. A., and Sedat, J. W. (1997). Interphase chromosomes undergo constrained diffusional motion in living cells. Curr. Biol. *7*, 930-939.
152. Matz, M. V., Fradkov, A. F., Labas, Y. A., Savitsky, A. P., Zaraisky, A. G., Markelov, M. L., and Lukyanov, S. A. (1999). Fluorescent proteins from nonbioluminescent Anthozoa species. Nat. Biotechnol. *17*, 969-973.
153. Maurer-Stroh, S., Dickens, N. J., Hughes-Davies, L., Kouzarides, T., Eisenhaber, F., and Ponting, C. P. (2003). The Tudor domain 'Royal Family': Tudor, plant Agenet, Chromo, PWWP and MBT domains. Trends Biochem. Sci. *28*, 69-74.
154. Maus, M., Rousseau, E., Cotlet, M., Schweitzer, G., Hofkens, J., Van der Auweraer, M., De Schryver, F. C., and Krueger, A. (2001). New picosecond laser system for easy tunability over the whole ultraviolet/visible/near infrared wavelength range based on flexible harmonic generation and optical parametric oscillation. Rev. Sci. Instrum. *72*, 36-40.
155. McDonald, D., Vodicka, M. A., Lucero, G., Svitkina, T. M., Borisy, G. G., Emerman, M., and Hope, T. J. (2002). Visualization of the intracellular behavior of HIV in living cells. J. Cell Biol. *159*, 441-452.
156. McKee, C. J., Kessl, J. J., Shkriabai, N., Dar, M. J., Engelman, A., and Kvaratskhelia, M. (2008). Dynamic modulation of HIV-1 integrase structure and function by cellular lens epithelium-derived growth factor (LEDGF) protein. J. Biol. Chem. *283*, 31802-31812.
157. Merzlyak, E. M., Goedhart, J., Shcherbo, D., Bulina, M. E., Shcheglov, A. S., Fradkov, A. F., Gaintzeva, A., Lukyanov, K. A., Lukyanov, S., Gadella, T. W., and Chudakov, D. M. (2007). Bright monomeric red fluorescent protein with an extended fluorescence lifetime. Nat. Methods *4*, 555-557.
158. Meyer, T. and Schindler, H. (1988). Particle counting by fluorescence correlation spectroscopy. Simultaneous measurement of aggregation and diffusion of molecules in solutions and in membranes. Biophys. J *54*, 983-993.
159. Meyvis, T. K., De Smedt, S. C., Van, O. P., and Demeester, J. (1999). Fluorescence recovery after photobleaching: a versatile tool for mobility and interaction measurements in pharmaceutical research. Pharm. Res. *16*, 1153-1162.
160. Michel, F., Crucifix, C., Granger, F., Eiler, S., Mouscadet, J. F., Korolev, S., Agapkina, J., Ziganshin, R., Gottikh, M., Nazabal, A., Emiliani, S., Benarous, R., Moras, D., Schultz, P., and Ruff, M. (2009). Structural basis for HIV-1 DNA integration in the human genome, role of the LEDGF/P75 cofactor. EMBO J. *28*, 980-991.
161. Michelman-Ribeiro, A., Mazza, D., Rosales, T., Stasevich, T. J., Boukari, H., Rishi, V., Vinson, C., Knutson, J. R., and McNally, J. G. (2009). Direct measurement of association and dissociation rates of DNA binding in live cells by fluorescence correlation spectroscopy. Biophys. J. *97*, 337-346.
162. Miyauchi, K., Kim, Y., Latinovic, O., Morozov, V., and Melikyan, G. B. (2009). HIV enters cells via endocytosis and dynamin-dependent fusion with endosomes. Cell *137*, 433-444.
163. Morerio, C., Acquila, M., Rosanda, C., Rapella, A., Tassano, E., Micalizzi, C., and Panarello, C. (2005). t(9;11)(p22;p15) with NUP98-LEDGF fusion gene in pediatric acute myeloid leukemia. Leuk. Res. *29*, 467-470.
164. Muchardt, C. and Yaniv, M. (1999). The mammalian SWI/SNF complex and the control of cell growth. Semin. Cell Dev. Biol. *10*, 189-195.
165. Mueller, F., Wach, P., and McNally, J. G. (2008). Evidence for a common mode of transcription factor interaction with chromatin as revealed by improved quantitative fluorescence recovery after photobleaching. Biophys. J. *94*, 3323-3339.
166. Muller, B. K., Zaychikov, E., Brauchle, C., and Lamb, D. C. (2005). Pulsed interleaved excitation. Biophys. J. *89*, 3508-3522.

167. Muller, C. B., Loman, A., Pacheco, V., Koberling, F., Willbold, D., Richtering, W., and Enderlein, J. (2008). Precise measurement of diffusion by multi-color dual-focus fluorescence correlation spectroscopy. Opt. Express *16*, 14353-14368.
168. Nakamura, H., Izumoto, Y., Kambe, H., Kuroda, T., Mori, T., Kawamura, K., Yamamoto, H., and Kishimoto, T. (1994). Molecular cloning of complementary DNA for a novel human hepatoma-derived growth factor. Its homology with high mobility group-1 protein. J. Biol. Chem. *269*, 25143-25149.
169. Nameki, N., Tochio, N., Koshiba, S., Inoue, M., Yabuki, T., Aoki, M., Seki, E., Matsuda, T., Fujikura, Y., Saito, M., Ikari, M., Watanabe, M., Terada, T., Shirouzu, M., Yoshida, M., Hirota, H., Tanaka, A., Hayashizaki, Y., Guntert, P., Kigawa, T., and Yokoyama, S. (2005). Solution structure of the PWWP domain of the hepatoma-derived growth factor family. Protein Sci. *14*, 756-764.
170. Nishizawa, Y., Usukura, J., Singh, D. P., Chylack, L. T., and Shinohara, T. (2001). Spatial and temporal dynamics of two alternatively spliced regulatory factors, lens epithelium-derived growth factor (ledgf/p75) and p52, in the nucleus. Cell and Tissue Research *305*, 107-114.
171. Nozaki, Y. (1972). The preparation of guanidine hydrochloride. Methods Enzymol. *26*, 43-50.
172. Ochs, R. L., Muro, Y., Si, Y., Ge, H., Chan, E. K., and Tan, E. M. (2000). Autoantibodies to DFS 70 kd/transcription coactivator p75 in atopic dermatitis and other conditions. J. Allergy Clin. Immunol. *105*, 1211-1220.
173. Oyama, R., Takashima, H., Yonezawa, M., Doi, N., Miyamoto-Sato, E., Kinjo, M., and Yanagawa, H. (2006). Protein-protein interaction analysis by C-terminally specific fluorescence labeling and fluorescence cross-correlation spectroscopy. Nucleic Acids Res. *34*, e102.
174. Paul, P. H., Garguilo, M. G., and Rakestraw, D. J. (1998). Imaging of pressure- and electrokinetically driven flows through open capillaries. Analytical Chemistry *70*, 2459-2467.
175. Petersen, N. O. (1986). Scanning fluorescence correlation spectroscopy. I. Theory and simulation of aggregation measurements. Biophys. J. *49*, 809-815.
176. Petersen, N. O., Hoddelius, P. L., Wiseman, P. W., Seger, O., and Magnusson, K. E. (1993). Quantitation of membrane receptor distributions by image correlation spectroscopy: concept and application. Biophys. J. *65*, 1135-1146.
177. Petrasek, Z. and Schwille, P. (2008). Precise measurement of diffusion coefficients using scanning fluorescence correlation spectroscopy. Biophys. J. *94*, 1437-1448.
178. Pettit, F. K., Bare, E., Tsai, A., and Bowie, J. U. (2007). HotPatch: a statistical approach to finding biologically relevant features on protein surfaces. J. Mol. Biol. *369*, 863-879.
179. Phair, R. D., Scaffidi, P., Elbi, C., Vecerova, J., Dey, A., Ozato, K., Brown, D. T., Hager, G., Bustin, M., and Misteli, T. (2004). Global nature of dynamic protein-chromatin interactions in vivo: three-dimensional genome scanning and dynamic interaction networks of chromatin proteins. Mol. Cell Biol. *24*, 6393-6402.
180. Podtelezhnikov, A. A., Gao, K., Bushman, F. D., and McCammon, J. A. (2003). Modeling HIV-1 integrase complexes based on their hydrodynamic properties. Biopolymers *68*, 110-120.
181. Poeschla, E. M. (2008). Integrase, LEDGF/p75 and HIV replication. Cell Mol. Life Sci. *65*, 1403-1424.
182. Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G., and Cormier, M. J. (1992). Primary structure of the Aequorea victoria green-fluorescent protein. Gene *111*, 229-233.

183. Prescott, M., Ling, M., Beddoe, T., Oakley, A. J., Dove, S., Hoegh-Guldberg, O., Devenish, R. J., and Rossjohn, J. (2003). The 2.2 A crystal structure of a pocilloporin pigment reveals a nonplanar chromophore conformation. Structure. *11*, 275-284.
184. Qiu, C., Sawada, K., Zhang, X., and Cheng, X. (2002a). The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. Nat. Struct. Biol. *9*, 217-224.
185. Qiu, C., Sawada, K., Zhang, X., and Cheng, X. (2002b). The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. Nat. Struct. Biol. *9*, 217-224.
186. Reeves, R. (2001). Molecular biology of HMGA proteins: hubs of nuclear function. Gene *277*, 63-81.
187. Ren, G., Gao, K., Bushman, F. D., and Yeager, M. (2007). Single-particle image reconstruction of a tetramer of HIV integrase bound to DNA. J. Mol. Biol. *366*, 286-294.
188. Rerks-Ngarm, S., Pitisuttithum, P., Nitayaphan, S., Kaewkungwal, J., Chiu, J., Paris, R., Premsri, N., Namwat, C., de, S. M., Adams, E., Benenson, M., Gurunathan, S., Tartaglia, J., McNeil, J. G., Francis, D. P., Stablein, D., Birx, D. L., Chunsuttiwat, S., Khamboonruang, C., Thongcharoen, P., Robb, M. L., Michael, N. L., Kunasol, P., and Kim, J. H. (2009). Vaccination with ALVAC and AIDSVAX to Prevent HIV-1 Infection in Thailand. N. Engl. J. Med. NEJMoa0908492.
189. Reznikoff, W. S. (2003). Tn5 as a model for understanding DNA transposition. Mol. Microbiol. *47*, 1199-1206.
190. Ries, J., Chiantia, S., and Schwille, P. (2009). Accurate determination of membrane dynamics with line-scan FCS. Biophys. J. *96*, 1999-2008.
191. Ries, J. and Schwille, P. (2006). Studying slow membrane dynamics with continuous wave scanning fluorescence correlation spectroscopy. Biophys. J. *91*, 1915-1924.
192. Rigler,R. and Elson,E. (2001). Fluorescence Correlation Spectroscopy - Theory and Applications. (Berlin: Springer-Verlag).
193. Rigler, R., Foldes-Papp, Z., Meyer-Almes, F. J., Sammet, C., Volcker, M., and Schnetz, A. (1998). Fluorescence cross-correlation: a new concept for polymerase chain reaction. J. Biotechnol. *63*, 97-109.
194. Ruan, Q., Cheng, M. A., Levi, M., Gratton, E., and Mantulin, W. W. (2004). Spatial-temporal studies of membrane dynamics: scanning fluorescence correlation spectroscopy (SFCS). Biophys. J. *87*, 1260-1267.
195. Saito, K., Wada, I., Tamura, M., and Kinjo, M. (2004). Direct detection of caspase-3 activation in single live cells by cross-correlation analysis. Biochem. Biophys. Res. Commun. *324*, 849-854.
196. Sali, A. and Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. *234*, 779-815.
197. Sarkar, I., Hauber, I., Hauber, J., and Buchholz, F. (2007). HIV-1 proviral DNA excision using an evolved recombinase. Science *316*, 1912-1915.
198. Saxton, M. J. (1994). Anomalous diffusion due to obstacles: a Monte Carlo study. Biophys. J. *66*, 394-401.
199. Saxton, M. J. (1996). Anomalous diffusion due to binding: a Monte Carlo study. Biophys. J. *70*, 1250-1262.
200. Schenk, A., Ivanchenko, S., Rocker, C., Wiedenmann, J., and Nienhaus, G. U. (2004a). Photodynamics of red fluorescent proteins studied by fluorescence correlation spectroscopy. Biophys. J. *86*, 384-394.
201. Schenk, A., Ivanchenko, S., Rocker, C., Wiedenmann, J., and Nienhaus, G. U. (2004b). Photodynamics of red fluorescent proteins studied by fluorescence correlation spectroscopy. Biophys. J. *86*, 384-394.

202. Schwille,P. (2001). Cross-correlation analysis in FCS. In Fluorescence Correlation Spectroscopy. Theory and Applications, E.L.Elson and R.Rigler, eds. (Berlin: Springer-Verlag), pp. 360-378.
203. Schwille,P. and Haustein,E. (2001). Fluorescence Correlation Spectroscopy: An Introduction to its Concepts and Applications. (Online: Biophysics Textbook Online).
204. Schwille, P., Korlach, J., and Webb, W. W. (1999). Fluorescence correlation spectroscopy with single-molecule sensitivity on cell and model membranes. Cytometry *36*, 176-182.
205. Schwille, P., Kummer, S., Heikal, A. A., Moerner, W. E., and Webb, W. W. (2000). Fluorescence correlation spectroscopy reveals fast optical excitation-driven intramolecular dynamics of yellow fluorescent proteins. Proc. Natl. Acad. Sci. U. S A *97*, 151-156.
206. Schwille, P., Meyer-Almes, F. J., and Rigler, R. (1997). Dual-color fluorescence cross-correlation spectroscopy for multicomponent diffusional analysis in solution. Biophys. J. *72*, 1878-1886.
207. Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N., Palmer, A. E., and Tsien, R. Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat. Biotechnol. *22*, 1567-1572.
208. Shaner, N. C., Lin, M. Z., McKeown, M. R., Steinbach, P. A., Hazelwood, K. L., Davidson, M. W., and Tsien, R. Y. (2008). Improving the photostability of bright monomeric orange and red fluorescent proteins. Nat. Methods *5*, 545-551.
209. Shaner, N. C., Steinbach, P. A., and Tsien, R. Y. (2005). A guide to choosing fluorescent proteins. Nat. Methods *2*, 905-909.
210. Sharma, P., Fatma, N., Kubo, E., Shinohara, T., Chylack, L. T., and Singh, D. P. (2003). Lens epithelium-derived growth factor relieves transforming growth factor-beta 1-induced transcription repression of heat shock proteins in human lens epithelial cells. Journal of Biological Chemistry *278*, 20037-20046.
211. Sharma, P., Singh, D. P., Fatma, N., Chylack, L. T., and Shinohara, T. (2000). Activation of LEDGF gene by thermal- and oxidative-stresses. Biochemical and Biophysical Research Communications *276*, 1320-1324.
212. Shi, X., Foo, Y. H., Sudhaharan, T., Chong, S. W., Korzh, V., Ahmed, S., and Wohland, T. (2009). Determination of dissociation constants in living zebrafish embryos with single wavelength fluorescence cross-correlation spectroscopy. Biophys. J *97*, 678-686.
213. Shimomura, O. (2005). The discovery of aequorin and green fluorescent protein. J. Microsc. *217*, 1-15.
214. Shinohara, T., Singh, D. P., and Fatma, N. (2002). LEDGF, a survival factor, activates stress-related genes. Prog. Retin. Eye Res. *21*, 341-358.
215. Shu, X., Shaner, N. C., Yarbrough, C. A., Tsien, R. Y., and Remington, S. J. (2006). Novel chromophores and buried charges control color in mFruits. Biochemistry *45*, 9639-9647.
216. Shun, M. C., Botbol, Y., Li, X., Di, N. F., Daigle, J. E., Yan, N., Lieberman, J., Lavigne, M., and Engelman, A. (2008). Identification and characterization of PWWP domain residues critical for LEDGF/p75 chromatin binding and human immunodeficiency virus type 1 infectivity. J. Virol. *82*, 11555-11567.
217. Shun, M. C., Raghavendra, N. K., Vandegraaff, N., Daigle, J. E., Hughes, S., Kellam, P., Cherepanov, P., and Engelman, A. (2007). LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. Genes Dev. *21*, 1767-1778.
218. Singh, D. P., Fatma, N., Kimura, A., Chylack, L. T., Jr., and Shinohara, T. (2001). LEDGF binds to heat shock and stress-related element to activate the expression of stress-related genes. Biochem. Biophys. Res. Commun. *283*, 943-955.
219. Singh, D. P., Kimura, A., Chylack, L. T., and Shinohara, T. (2000a). Lens epithelium-derived growth factor (LEDGF/p75) and p52 are derived from a single gene by alternative splicing. Gene *242*, 265-273.

220. Singh, D. P., Kubo, E., Takamura, Y., Shinohara, T., Kumar, A., Chylack, L. T., Jr., and Fatma, N. (2006). DNA binding domains and nuclear localization signal of LEDGF: contribution of two helix-turn-helix (HTH)-like domains and a stretch of 58 amino acids of the N-terminal to the trans-activation potential of LEDGF. J. Mol. Biol. *355*, 379-394.
221. Singh, D. P., Ohguro, N., Kikuchi, T., Chylack, L. T., and Shinohara, T. (1998). Gene sequence and functional studies of lens epithelial cell derived growth factor (LEDGF). Investigative Ophthalmology & Visual Science *33*, S777.
222. Singh, D. P., Ohguro, N., Kikuchi, T., Sueno, T., Reddy, V. N., Yuge, K., Chylack, L. T., and Shinohara, T. (2000b). Lens epithelium-derived growth factor: Effects on growth and survival of lens epithelial cells, keratinocytes, and fibroblasts. Biochemical and Biophysical Research Communications *267*, 373-381.
223. Soumpasis, D. M. (1983). Theoretical analysis of fluorescence photobleaching recovery experiments. Biophys. J. *41*, 95-97.
224. Sprague, B. L. and McNally, J. G. (2005). FRAP analysis of binding: proper and fitting. Trends Cell Biol. *15*, 84-91.
225. Sprague, B. L., Pego, R. L., Stavreva, D. A., and McNally, J. G. (2004). Analysis of binding reactions by fluorescence recovery after photobleaching. Biophys. J. *86*, 3473-3495.
226. Stec, I., Nagl, S. B., van Ommen, G. J., and den Dunnen, J. T. (2000). The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? FEBS Lett. *473*, 1-5.
227. Stepanenko, O. V., Verkhusha, V. V., Kazakov, V. I., Shavlovsky, M. M., Kuznetsova, I. M., Uversky, V. N., and Turoverov, K. K. (2004). Comparative studies on the structure and stability of fluorescent proteins EGFP, zFP506, mRFP1, "dimer2", and DsRed1. Biochemistry *43*, 14913-14923.
228. Sue, S. C., Chen, J. Y., Lee, S. C., Wu, W. G., and Huang, T. H. (2004). Solution structure and heparin interaction of human hepatoma-derived growth factor. J. Mol. Biol. *343*, 1365-1377.
229. Sue, S. C., Lee, W. T., Tien, S. C., Lee, S. C., Yu, J. G., Wu, W. J., Wu, W. G., and Huang, T. H. (2007). PWWP module of human hepatoma-derived growth factor forms a domain-swapped dimer with much higher affinity for heparin. J. Mol. Biol. *367*, 456-472.
230. Sugaya, K., Vigneron, M., and Cook, P. R. (2000). Mammalian cell lines expressing functional RNA polymerase II tagged with the green fluorescent protein. J. Cell Sci. *113 (Pt 15)*, 2679-2683.
231. Sykora, J., Kaiser, K., Gregor, I., Bonigk, W., Schmalzing, G., and Enderlein, J. (2007). Exploring fluorescence antibunching in solution to determine the stoichiometry of molecular complexes. Anal. Chem. *79*, 4040-4049.
232. Takahashi, Y., Nishimura, J., Suzuki, A., Ishibashi, K., Kinjo, M., and Miyawaki, A. (2008). Cross-talk-free fluorescence cross-correlation spectroscopy by the switching method. Cell Struct. Funct. *33*, 143-150.
233. Thakar, K., Niedenthal, R., Okaz, E., Franken, S., Jakobs, A., Gupta, S., Kelm, S., and Dietz, F. (2008). SUMOylation of the hepatoma-derived growth factor negatively influences its binding to chromatin. FEBS J. *275*, 1411-1426.
234. Thews, E., Gerken, M., Eckert, R., Zapfel, J., Tietz, C., and Wrachtrup, J. (2005). Cross talk free fluorescence cross correlation spectroscopy in live cells. Biophys. J. *89*, 2069-2076.
235. Thompson, N. L., Burghardt, T. P., and Axelrod, D. (1981). Measuring surface dynamics of biomolecules by total internal reflection fluorescence with photobleaching recovery or correlation spectroscopy. Biophys. J. *33*, 435-454.
236. Tinoco,I., Sauer,K., Wang,J.S., and Puglisi,J.D. (2001). Physical Chemistry: Principles and Applications in Biological Sciences. (Upper Saddle River, NJ: Prentice Hall).

237. Tsiang, M., Jones, G. S., Hung, M., Mukund, S., Han, B., Liu, X., Babaoglu, K., Lansdon, E., Chen, X., Todd, J., Cai, T., Pagratis, N., Sakowicz, R., and Geleziunas, R. (2009). Affinities between the binding partners of the HIV-1 integrase dimer-lens epithelium-derived growth factor (IN dimer-LEDGF) complex. J. Biol. Chem. *284*, 33580-33599.
238. Turlure, F., Maertens, G., Rahman, S., Cherepanov, P., and Engelman, A. (2006). A tripartite DNA-binding element, comprised of the nuclear localization signal and two AT-hook motifs, mediates the association of LEDGF/p75 with chromatin in vivo. Nucleic Acids Res. *34*, 1653-1675.
239. Van Craenenbroek, E., Vercammen, J., Matthys, G., Beirlant, J., Marot, C., Hoebeke, J., Strobbe, R., and Engelborghs, Y. (2001). Heuristic statistical analysis of fluorescence fluctuation data with bright spikes: application to ligand binding to the human serotonin receptor expressed in Escherichia coli cells. Biol. Chem. *382*, 355-361.
240. Van Maele, B. and Debyser, Z. (2005). HIV-1 integration: an interplay between HIV-1 integrase, cellular and viral proteins. AIDS Rev. *7*, 26-43.
241. Vandekerckhove, L., Christ, F., Van Maele, B., De Rijck, J., Gijsbers, R., Van den Haute, C., Witvrouw, M., and Debyser, Z. (2006). Transient and stable knockdown of the integrase cofactor LEDGF/p75 reveals its role in the replication cycle of human immunodeficiency virus. Journal of Virology *80*, 1886-1896.
242. Vanegas, M., Llano, M., Delgado, S., Thompson, D., Peretz, M., and Poeschla, E. (2005). Identification of the LEDGF/p75 HIV-1 integrase-interaction domain and NLS reveals NLS-independent chromatin tethering. J. Cell Sci. *118*, 1733-1743.
243. Vercammen, J., Maertens, G., Gerard, M., De, C. E., Debyser, Z., and Engelborghs, Y. (2002). DNA-induced polymerization of HIV-1 integrase analyzed with fluorescence fluctuation spectroscopy. J. Biol. Chem. *277*, 38045-38052.
244. Verkhusha, V. V. and Sorkin, A. (2005b). Conversion of the monomeric red fluorescent protein into a photoactivatable probe. Chem. Biol. *12*, 279-285.
245. Verkhusha, V. V. and Sorkin, A. (2005a). Conversion of the monomeric red fluorescent protein into a photoactivatable probe. Chem. Biol. *12*, 279-285.
246. Violot, S., Hong, S. S., Rakotobe, D., Petit, C., Gay, B., Moreau, K., Billaud, G., Priet, S., Sire, J., Schwartz, O., Mouscadet, J. F., and Boulanger, P. (2003). The human polycomb group EED protein interacts with the integrase of human immunodeficiency virus type 1. J. Virol. *77*, 12507-12522.
247. Vogel, H. (1921). Das temperature-abhängigketsgesetz der viskosität von flüssigkeiten. Physikalische Zeitschrift *22*, 645.
248. Vosch, T., Cotlet, M., Hofkens, J., Van der Biest, K., Lor, M., Weston, K., Tinnefeld, P., Sauer, M., Latterini, L., Mullen, K., and De Schryver, F. C. (2003). Probing Forster type energy pathways in a first generation rigid dendrimer bearing two perylene imide chromophores. J. Phys. Chem. A *107*, 6920-6931.
249. Wachsmuth, M., Waldeck, W., and Langowski, J. (2000). Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy. J. Mol. Biol. *298*, 677-689.
250. Wang, J. Y., Ling, H., Yang, W., and Craigie, R. (2001). Structure of a two-domain fragment of HIV-1 integrase: implications for domain organization in the intact protein. EMBO J. *20*, 7333-7343.
251. Wang, L., Jackson, W. C., Steinbach, P. A., and Tsien, R. Y. (2004). Evolution of new nonantibody proteins via iterative somatic hypermutation. Proc. Natl. Acad. Sci. U. S. A *101*, 16745-16749.
252. Wawrezinieck, L., Rigneault, H., Marguet, D., and Lenne, P. F. (2005). Fluorescence correlation spectroscopy diffusion laws to probe the submicron cell membrane organization. Biophys. J. *89*, 4029-4042.

253. Weber, W., Helms, V., McCammon, J. A., and Langhoff, P. W. (1999). Shedding light on the dark and weakly fluorescent states of green fluorescent proteins. Proc. Natl. Acad. Sci. U. S. A *96*, 6177-6182.
254. Wedemeier, A., Zhang, T., Merlitz, H., Wu, C. X., and Langowski, J. (2008). The role of chromatin conformations in diffusional transport of chromatin-binding proteins: Cartesian lattice simulations. J. Chem. Phys. *128*, 155101.
255. Weiss, M., Elsner, M., Kartberg, F., and Nilsson, T. (2004). Anomalous subdiffusion is a measure for cytoplasmic crowding in living cells. Biophys. J. *87*, 3518-3524.
256. Widengren, J., Mets, U., and Rigler, R. (1995). Fluorescence Correlation Spectroscopy of Triplet-States in Solution - A Theoretical and Experimental-Study. Journal of Physical Chemistry *99*, 13368-13379.
257. Widengren, J., Mets, U., and Rigler, R. (1999b). Photodynamic properties of green fluorescent proteins investigated by fluorescence correlation spectroscopy. Chem. Phys. *250*, 171-186.
258. Widengren, J., Mets, U., and Rigler, R. (1999a). Photodynamic properties of green fluorescent proteins investigated by fluorescence correlation spectroscopy. Chemical Physics *250*, 171-186.
259. Wiedenmann, J., Schenk, A., Rocker, C., Girod, A., Spindler, K. D., and Nienhaus, G. U. (2002). A far-red fluorescent protein with fast maturation and reduced oligomerization tendency from Entacmaea quadricolor (Anthozoa, Actinaria). Proc. Natl. Acad. Sci. U. S. A *99*, 11646-11651.
260. Wiseman, P. W., Squier, J. A., Ellisman, M. H., and Wilson, K. R. (2000). Two-photon image correlation spectroscopy and image cross-correlation spectroscopy. J. Microsc. *200*, 14-25.
261. Wu, B. and Muller, J. D. (2005). Time-integrated fluorescence cumulant analysis in fluorescence fluctuation spectroscopy. Biophys. J. *89*, 2721-2735.
262. Wu, X., Daniels, T., Molinaro, C., Lilly, M. B., and Casiano, C. A. (2002). Caspase cleavage of the nuclear autoantigen LEDGF/p75 abrogates its pro-survival function: implications for autoimmunity in atopic disorders. Cell Death and Differentiation *9*, 915-925.
263. Yarbrough, D., Wachter, R. M., Kallio, K., Matz, M. V., and Remington, S. J. (2001). Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-A resolution. Proc. Natl. Acad. Sci. U. S. A *98*, 462-467.
264. Yokoyama, A. and Cleary, M. L. (2008). Menin critically links MLL proteins with LEDGF on cancer-associated target genes. Cancer Cell *14*, 36-46.
265. Yung, E., Sorin, M., Pal, A., Craig, E., Morozov, A., Delattre, O., Kappes, J., Ott, D., and Kalpana, G. V. (2001). Inhibition of HIV-1 virion production by a transdominant mutant of integrase interactor 1. Nat. Med. *7*, 920-926.
266. Yung, E., Sorin, M., Wang, E. J., Perumal, S., Ott, D., and Kalpana, G. V. (2004). Specificity of interaction of INI1/hSNF5 with retroviral integrases and its functional significance. J. Virol. *78*, 2222-2231.
267. Zheng, R., Jenkins, T. M., and Craigie, R. (1996). Zinc folds the N-terminal domain of HIV-1 integrase, promotes multimerization, and enhances catalytic activity. Proc. Natl. Acad. Sci. U. S. A *93*, 13659-13664.
268. Zielske, S. P. and Stevenson, M. (2006). Modest but reproducible inhibition of human immunodeficiency virus type 1 infection in macrophages following LEDGFp75 silencing. J. Virol. *80*, 7275-7280.
269. Zlatanova, J., Caiafa, P., and Van, H. K. (2000). Linker histone binding and displacement: versatile mechanism for transcriptional regulation. FASEB J. *14*, 1697-1704.

Curriculum vitae

Personalia

Name	Jelle Hendrix
Address	grotestraat 112 3540 Herk-de-Stad
E-mail address	jelle_hendrix@hotmail.com jelle.hendrix@chem.kuleuven.be
Telephone	+32 (0) 473 842024
Day/Place of birth	April 28, 1983, Hasselt

Higher education

2001-2003	Kandidate in Chemistry (KUL)
2003-2005	Licentiate (Master) in Biochemistry (Magna cum laude) (KUL) Thesis: Study of the interaction of HIV-1 integrase and LEDGF/p75 and characterization of monomeric red fluorescent proteins with fluorescence correlation spectroscopy
2005-2010	PhD in Science, option Biochemistry Scholarship from the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT)

List of publications

Bartholomeeusen, K., Christ, F., Hendrix, J., Rain, J.C., Emiliani, S., Benarous, R., Debyser, Z., Gijsbers, R., and De Rijck, J. (2009). Lens Epithelium Derived Growth Factor/p75 interacts with the transposase derived DDE domain of pogZ. J. Biol. Chem. 284, 11467-11477.

Briers, Y., Schmelcher, M., Loessner, M.J., Hendrix, J., Engelborghs, Y., Volckaert, G., and Lavigne, R. (2009). The high-affinity peptidoglycan binding domain of Pseudomonas phage endolysin KZ144. Biochem. Biophys. Res. Commun. 383, 187-191.

Buyens, K., Lucas, B., Raemdonck, K., Braeckmans, K., Vercammen, J., Hendrix, J., Engelborghs, Y., De Smedt, S.C., and Sanders, N.N. (2008). A fast and sensitive method for measuring the integrity of siRNA-carrier complexes in full human serum. J. Control Release 126, 67-76.

Daniëls, V., Vancraenenbroeck, R., Law, B. M., Greggio, E., Hendrix, J., Lobbestael, E., Gao, F., De Maeyer, M., Cookson, M. R., Harvey, K., Baekelandt, V., and Taymans, J. M. The interaction between ROC GTPase and COR domains of Leucine-rich repeat kinase 2 (LRRK2) as a regulatory mechanism of LRRK2 GTPase activity. (submitted for publication)

De Graeve, S., Marinelli, S., Stolz, F., Hendrix, J., Engelborghs, Y., Van Dijck, P., and Thevelein, J.M. (2009). Mammalian brain proteins that counteract or enhance alpha-synuclein toxicity in yeast. (submitted for publication)

De Rijck, J., Vandekerckhove, L., Gijsbers, R., Hombrouck, A., Hendrix, J., Vercammen, J., Engelborghs, Y., Christ, F., and Debyser, Z. (2006). Overexpression of the lens epithelium-derived growth factor/p75 integrase binding domain inhibits human immunodeficiency virus replication. J. Virol. 80, 11498-11509.

Hendrix, J., De Rijck, J., Gijsbers, R., Voet, A., Hotta, J., McNeely, M., Vanstreels, E., Daelemans, D., Hofkens, J., Debyser, Z., and Engelborghs, Y. Dynamic chromatin scanning by transcriptional co-activator LEDGF/p75 is arrested by HIV-1 integrase. (submitted for publication)

Hendrix, J., Engelborghs, Y., Quantitative fluorescence cross-correlation spectroscopy with fluorescent proteins. (manuscript in preparation)

Hendrix, J., Flors, C., Dedecker, P., Hofkens, J., and Engelborghs, Y. (2008). Dark states in monomeric red fluorescent proteins studied by fluorescence correlation and single molecule spectroscopy. Biophys. J. 94, 4103-4113.
Hombrouck, A., De Rijck, J., Hendrix, J., Vandekerckhove, L., Voet, A., De Maeyer, M., Witvrouw, M., Engelborghs, Y., Christ, F., Gijsbers, R., and Debyser, Z. (2007). Virus Evolution Reveals an Exclusive Role for LEDGF/p75 in Chromosomal Tethering of HIV. PLoS. Pathog. 3, e47.

Ricicová, M., Tournu, H., Kucharíková, S., Hendrix, J., Bujdáková, H., Van Eldere, J., Lagrou, K., and Van Dijck, P. (2009). Candida albicans biofilm development in a simple in vivo subcutaneous rat model is similar to the central venous catheter model. (submitted for publication)

Lectures given at international conferences

Investigation of the HIV-1 integrase interactor LEDGF/p75 in human cells with Fluorescence Correlation and Cross-Correlation Spectroscopy

Targeting replication and integration of HIV - TRIoH General Assembly

Barcelona, Spain, 8-12/12/06

Dark States in Monomeric Red Fluorescent Proteins studied by FCS and Single Molecule Spectroscopy

Joint Biophysical Society 52nd Annual Meeting and 16th IUPAB International Biophysics Congress

Long Beach, California, USA, 2-6/02/08

LEDGF/p75 switches from a dynamic to a tight chromatin interaction upon binding to HIV-1 integrase

Targeting HIV integration co-factors - 2nd General Assembly

Prague, Czech Republic, 29-30/01/09

LEDGF/p75 switches from a dynamic to a tight chromatin interaction upon binding to HIV-1 integrase

European Biophysics Congress 2009

Genoa, Italy, 11-15/07/09

Cellular tunable focus FCS for studying the molecular mechanism of LEDGF/p75 mediated chromatin tethering of HIV-1 integrase

12th Carl Zeiss sponsored workshop on FCS and related methods

Cargèse, Corsica, 12-16/10/09

Posters presented at international conferences

Investigation of the HIV-1 integrase interactor LEDGF/p75 in human cells with Fluorescence Correlation and Cross-Correlation Spectroscopy

Targeting replication and integration of HIV - TRIoH General Assembly

Barcelona, Spain, 8-12/12/06

Investigation of the HIV-1 integrase interactor LEDGF/p75 in human cells with Fluorescence Correlation and Cross-Correlation Spectroscopy

Zeiss International Workshop on Fluorescence Correlation Spectroscopy methods

Stockholm, Sweden, 4-6/12/06

How a human protein becomes a hostage of the HIV - Fluorescence techniques for studying live-cell protein dynamics reveal strong chromatin tethering of HIV-1 integrase by human LEDGF/p75

Joint Biophysical Society 52nd Annual Meeting and 16th IUPAB International Biophysics Congress

Long Beach, California, USA, 2-6/02/08

1. How a human protein becomes a hostage of the HIV - Fluorescence techniques for studying live-cell protein dynamics reveal strong chromatin tethering of HIV-1 integrase by human LEDGF/p75
2. Dark States in Monomeric Red Fluorescent Proteins studied by Fluorescence Correlation and Single Molecule Spectroscopy

7th International Weber Symposium on Innovative Fluorescence Methodologies in Biochemistry and Medicine and 11th International Workshop on Fluorescence Correlation Spectroscopy and Related Methods

Kauai, Hawaii, USA, 6-12/06/08

Awards

LEDGF/p75 switches from a dynamic to a tight chromatin interaction upon binding to HIV-1 integrase

Targeting HIV integration co-factors, 2nd General Assembly

Prague, Czech Republic, 29-30/01/09

Best poster presentation

LEDGF/p75 switches from a dynamic to a tight chromatin interaction upon binding to HIV-1 integrase

Belgian Physical Society & Belgian Biophysical Society General Scientific Meeting

Hasselt, 01/04/09

European Physical Journal award for best poster presentation